Constraints in applied mathematics : rods, membranes, and cuckoos

Constraints in applied mathematics

Rods, membranes, and cuckoos

Graphics on the cover: the Weaire-Phelan foam is a counterexample to Kelvin's conjecture about the best partition of space into equal-volume cells. Reprinted with permission. Copyright John M. Sullivan.

Constraints in applied mathematics

Rods, membranes, and cuckoos

Proefschrift

ter verkrijging van de graad van Doctor
aan de Technische Universiteit Eindhoven,
op gezag van de Rector Magnificus, prof.dr.ir. C.J. van Duijn,
voor een commissie aangewezen door het
College voor Promoties in het openbaar te verdedigen
op donderdag 7 april 2005 om 16.00 uur

door

Robert Planqué

geboren te Leiderdorp

Dit proefschrift is goedgekeurd door de promotor:

prof.dr. M.A. Peletier

Copromotor:
dr. G.H.M. van der Heijden

Het onderzoek dat tot dit proefschrift heeft geleid werd mede mogelijk gemaakt door het Centrum voor Wiskunde en Informatica en het Thomas Stieltjes Institute for Mathematics.

THOMAS STIELTJES INSTITUTE
FOR MATHEMATICS

Toda lógica contém inevitável dose de mistificação. Toda mistificação contém boa dose de inevitável verdade, precisamos também do obscuro.
All logic contains an inevitable dose of mystique. All mystique contains a large dose of inevitable truth, we need the obscure as well.

———João Guimaraes Rosa

Sir Edmund Blackadder: "I see, and the fact that this secret has eluded the most intelligent people since the dawn of time doesn't at all dampen your spirits?"
Percy: "Oh no; I like a challenge!"

———Blackadder II:4

Contents

Preface

This thesis is the result of an accumulation of mathematical results compiled in the four years between December 2000 and December 2004 at the Centrum voor Wiskunde en Informatica in Amsterdam. Some preliminary studies described in the last chapter were conducted whilst in Bath, UK, where I was an exchange student with Nick Britton and Nigel Franks for some months.

This thesis bears the name of a single author, but at its basis lies the close collaboration I have enjoyed first and foremost with my promotor Mark Peletier. Each of the many meetings we have had over the years has added a valuable lesson to all parts of research, from 'how to hang a whiteboard', to 'how to write more thorough proofs'. I will especially remember the times at Mark and Irene's homes in Amsterdam and later Eindhoven, in which many incremental results were obtained in the finest way of doing science—armed with chocolate and coffee, cracking jokes all afternoon, and having a guitar at hand when one is running out of ideas or when the brain simply runs out of steam.

I also wish to express my gratitude for the smoothness of collaboration with my copromotor Gert van der Heijden, and for his invitation for Mark and I to visit him at UCL in London. I have felt most welcome there, and have greatly enjoyed our extensive discussions in front of the whiteboard at the Centre for Nonlinear Dynamics and its Applications.

It was Bert Peletier's enthusiasm and encouragement for my idea to go abroad during my undergraduate studies that provided the final push to take it forward. I would like to thank him for that support and for his suggestion to team up with Mark at the CWI.

My stay in Bath has led to more than I would have ever dreamt of five years ago. What started as a student's project on those enigmatic cuckoos has grown into a full-fledged collaboration with Nick Britton, Nigel Franks, and Tomáš Grim, and from there to my current job. I hope we will continue our joint efforts for years to come.

I wish to thank all my colleagues at the CWI, for making the place the fine work environment it is. In particular, my room mates: Carlota and Ignacio for helping me get into the maths and for keeping up my Spanish, Carolynne for her cheerfulness at 8:30 in the morning, Jeroen for sharing the fun building those paper planes, JF for his wit and endless matlab magic, and Maksat for filling me in on a country I knew nothing about; and 'the guys next door': Danijela for holding up a mirror by showing me how things can be done entirely differently, Bernard for his steady supply of cookies, Andrea for all the laughs and dinners, and Maciej for his many good tips and sense of humour.

And finally, of course, my family, who have endured many an enthusiastic but probably rather unintelligible tale about all the research, and have always been there to offer advice and insight. This is naturally most true for Manu, who has lovingly stood by me over these years. She has shared all the ups and downs that make up the life of a researcher. I hope we will share many more.

Bristol, February 21, 2005.

CHAPTER 1

Introduction

This thesis consists of two parts: constrained energy minimization and evolutionary ecology.

We start with a short primer on constrained energy minimization in Section 1.1. Once the reader is familiar with some of its basic concepts, two problems attacked in Chapters 2 through 5 are introduced: Sections 1.2, 1.3, and 1.4 lay the foundations to understand how an elastic rod buckles around a cylinder; the other minimization problem we study, the modeling of lipid bilayers, is introduced in Section 1.5.

An evolutionary ecological problem on how host birds should defend themselves against a parasitic cuckoo, is kept for the final Section 1.6 of this introduction.

1.1. Energy minimization and contact problems

The field of optimization has applications in many branches of science and technology. Economists use optimization techniques to maximize profits given a set of resources; engineers employ them to optimally control manufacturing processes to ensure the desired quality of the product; experimental scientists of all breeds benefit from them to generate the best approximation from measurements of their object under scrutiny. The goal in each of these applications is immediately clear. The problems under investigation in this thesis, however, are physical problems, and we may wonder what use optimization techniques have here. To motivate the value of the minimization viewpoint in this case, we commence by sketching the rationale behind the approach specifically for problems from physics.

Nature does not work in arbitrary ways. We have found many laws by which the constituents of our world abide, ranging from the theories of relativity and quantum mechanics to evolution. In many cases,

such laws are the manifestation of a quest for economics, and these are the topic of this introduction. Take, for instance, a soap bubble. This bubble does not have just any shape: when floating through the air it seems perfectly spherical. Such optimal geometry may be found in many places, such as the mirror-like surface of a lake on a wind still day, or the parabolic trajectory of a bullet.

In fact, we can understand a great deal of behaviour in our own society by viewing it as minimizations of effort and costs: we take the shortest route to work to save time, look for the best deals at the supermarket to save money, or arrange our daily tasks in such a way to obtain maximal efficiency, to name just a few examples.

The actual behaviour of a system in the natural world can thus often be understood by viewing it as the minimal behaviour, in a particular specified sense, within a range of possibilities. For example, if one takes a piece of string and holds it with both hands at the ends at the same height, the string hangs down with the lowest part in the middle between the hands. There are a multitude of possible configurations for this string while keeping it fixed at both ends, but its actual shape is a regular, symmetric one. The unifying principle at work in such situations is that we can formulate a certain energy for the system under consideration, such that the characteristics of the system—its dynamics, or its form—are those that belong to the configuration that has the least energy among all possible configurations.

This principle forms our starting point for the mathematical enquiries in Chapters 2, 4 and 5.

Taking constraints into account

Consider the problem of a rigid body being thrown from some location. Newton's second law of motion,

$$F = m\frac{d^2x}{dt^2},$$

provides us with a complete solution of the location of the rigid body that can be modelled as a point mass: given the force F that is exerted on the rigid body, and the mass m of the object, we know its acceleration, and hence its subsequent position x by integration. If the idealization of a point mass cannot be admitted, the problem immediately becomes very cumbersome: we then have to take all internal forces into account that result in the rigidity of the body to evaluate the resulting force. This is often a daunting task.

The method of energy minimization provides an elegant and very flexible way of dealing with such situations. Let us go through an example by asking ourselves the following simple question: why is the level of the water surface for a glass of water horizontal? We will show that, among all possible configurations with a specified amount of water in the glass, the configuration with a flat surface has lowest energy. Let us first assume that the water is at rest, i.e., the water is not moving. Let the surface forming the bottom of the glass be denoted by Ω, and let $u \in L^1(\Omega; \mathbb{R}^+)$ denote the height of the water level above some reference height. The potential energy of the glass is given by

$$F(u) = \frac{1}{2}\rho g \int_\Omega u^2(x)dx,$$

where ρ is the density of water and g is the gravitational constant at sea level. Now, instead of describing in some geometric fashion how a certain amount of water may be positioned within the glass, we merely constrain the set of functions by specifying the amount of water. This is given by

$$J(u) = \int_\Omega u(x)dx.$$

We now want to minimize the $F(u)$ over the set of all admissible function, which we will denote by K_m. This set is simply given by

$$K_m = \{u \in L^1 \mid J(u) = m\}$$

for any $m > 0$. So let us consider the minimization problem

$$\min\{F(u) \mid u \in K_m\}.$$

Let $v \in L^1$ be such that $\int_\Omega v(x)dx = 0$. Then $u + \varepsilon v$ is again an element of K_m for any $\varepsilon > 0$, and u is a minimizer if

$$F(u) \leq F(u + \varepsilon v) \quad \text{for all } v \text{ such that } \int_\Omega v = 0.$$

Hence $F'(u) \cdot v \geq 0$ for all such v. We compute

$$F'(u) \cdot v = \rho g \int_\Omega uv, \quad J'(u) \cdot v = \int_\Omega v.$$

If u is a minimizer, then there exists a $\lambda \in \mathbb{R}$ such that

$$F'(u) = \lambda J'(u). \tag{1.1}$$

This equation is called the Euler-Lagrange equation belonging to the minimization problem. Euler-Lagrange equations form an important

part of the study of minimization problems. The graphical interpretation of (1.1) is shown in Figure 1.1. So, absorbing the ρg into λ, we obtain

$$\int_\Omega uv = \lambda \int_\Omega v \quad \text{for all } v \text{ such that } \int_\Omega v = 0.$$

We conclude $\int_\Omega v(u - \lambda) = 0$ for all v such that $\int_\Omega v = 0$, and therefore $u = \lambda$ almost everywhere. Putting this back into context, we have derived that the water surface in a glass of water is indeed horizontal.

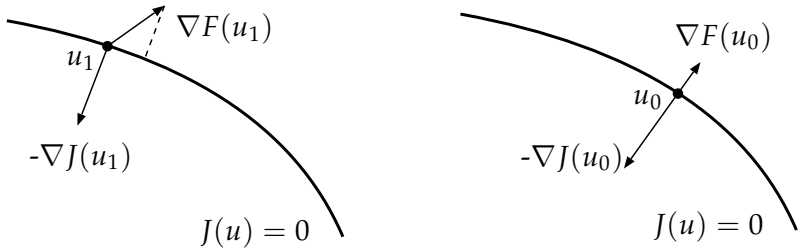


Figure 1.1: Graphical representation of the Lagrange multiplier method. If $\{F(u) \mid J(u) = 0\}$ is not minimized at u_1 (left), then $\nabla F(u_1)$ still has a non-zero projection (dashed line) onto the tangent at u_1, indicating we can improve on $F(u_1)$. If $\{F(u) \mid J(u) = 0\}$ is minimized at u_0 (right), then the derivatives $\nabla J(x_0)$ and $\nabla F(x_0)$ are aligned, and hence are scalar multiples. This constant is the Lagrange multiplier.

Problem setting

The previous example gives a good picture of the mathematical setting of the minimization problems we want to study. First we review some concepts. Let U, V be normed vector spaces. A set $P \subset U$ is called a *cone* if $p \in P$ implies $\lambda p \in P$ for all $\lambda \in \mathbb{R}^+$. For $u, v \in V$ we write $u \geq v$ (with respect to P) if $u - v \in P$. The cone P defining this relation is called the *positive cone* in V.

For the purposes of this thesis, the central problem of minimizing a constrained energy functional is now stated as follows:

Let U and V be normed vector spaces. Let $F : U \to \mathbb{R}$ be a functional and $J : U \to V$ be a constraint operator, and let $P \subset V$ be a positive cone. Find an element $\bar{u} \in U$ such that

$$F(\bar{u}) \leq F(u) \quad \forall u \in U \text{ such that } J(u) \in P. \tag{1.2}$$

The most important example here is the cone generated by a linear (or affine) inequality constraint $J(u) \geq 0$, such as

$$J(u) = u \geq 0, \quad \text{or} \quad J(u) = \left(m + \int u, m - \int u \right) \geq (0, 0).$$

In this thesis we study minimization problems that are simultaneously subjected to multiple constraints.

Existence: direct methods

The first important task upon defining a minimization problem is to show that it is well-defined. That is, to prove that indeed a minimizer exists within the class of admissible functions. The following example shows that even in the simplest cases there may not exist any minimizer. Consider the minimization problem:

$$\inf\{x \in (0, 1)\}. \tag{1.3}$$

It is evident that there does not exist any number $\bar{x} \in (0, 1)$ such that $\bar{x} \leq x$ for all $x \in (0, 1)$. Of course, if we augment the set of admissible numbers by adding 0, we do obtain a well-posed problem.

We mainly follow the so-called 'direct methods' (for an overview of such methods, see [32]). Consider the minimization problem

$$\inf\{F(u) \mid u \in U\}.$$

To prove existence of a minimizer the following method is used. Choose a minimizing sequence $\{u_n\}$ of functions and prove that it contains a subsequence that converges in some manner to a limit function $\bar{u}$ that lies in U. Moreover, prove that $\bar{u}$ is indeed a minimizer, i.e., that F is lower semicontinuous:

$$F(\bar{u}) \leq \liminf_{n \to \infty} F(u_n), \text{ if } u_n \text{ converges to } \bar{u}.$$

There is an inevitable tension between these two requirements: to prove convergence of a subsequence, and for the limit to remain in U, we wish to have as weak a topology on U; however, the stronger the topology, the easier lower semicontinuity is satisfied.

In practise, the minimizing sequence often suffers from a lack of compactness in the canonical topology of U, the norm topology. In (1.3), for example, if we take a minimizing sequence, it is clear that its limit will be 0, which is not a member of the set. It is the lack of compactness of the sequence that causes the ill-posedness of the problem here.

In this thesis we find a similar obstacle caused by a possible lack of compactness. To prove compactness of a minimizing sequence in Chapter 5, we set $U = L^1(\mathbb{R})$. This *a priori* gives very little regularity of the minimizer, but simply minimizing over a more restrictive class (such as $H^1(\mathbb{R})$, which implies continuity for a start) would not help: although we have more control over the form of the functions over which we minimize, it becomes more difficult to prove that the limit function $\bar{u}$ is still in H^1. To prove this we would for instance have to prove *independently* that $\bar{u}$ is continuous, without having the Euler-Lagrange equations at our disposal.

Energy minimization problems now have two main objects of study:

- Stationary points (local minimizers), which are attacked with the Euler-Lagrange equations.
- Global minimizers, which are scrutinized using cut and paste techniques.

Study of stationary points: Euler-Lagrange equations

Any minimizer is also a stationary point, i.e., a function for which the derivative of the functional at this function vanishes in every admissible direction. These stationary points are thus by definition defined as the solutions of the Euler-Lagrange equations that correspond to the minimization problem.

The Euler-Lagrange equations for constrained problems such as in (1.2) have been derived for very general settings. A comprehensive overview on this subject can be found in [91].

We briefly review the most important theorem regarding minimization problems involving inequality constraints, the Generalized Kuhn-Tucker Theorem [91, p. 249]. The precise conditions under which this theorem holds are stated below, but one important definition is not dealt with here, that of a regular point. The precise statement of that concept is rather involved and beyond the scope of this thesis: in Chapters 2 and 5 we prove our own versions of this theorem, with self-contained proofs.

Let U be a normed vector space. A functional $F : U \to \mathbb{R}$ is called *Fréchet differentiable* in $u \in U$ if for all $v \in U$, there exists a map $T(u;\cdot) : U \to \mathbb{R}$ which is linear and continuous in its argument, such

that

$$\lim_{\|v\|_U \to 0} \frac{\|F(u+v) - F(u) - T(u;v)\|_U}{\|v\|_U} = 0.$$

The linear map $T(u;v)$ is denoted by $F'(u) \cdot v$. We denote the dual of a U by U', and the pairing between elements in U and U' by ${}_U\langle \cdot, \cdot \rangle_{U'}$.

THEOREM 1.1 (Generalized Kuhn-Tucker Theorem). *Let U and V be normed vector spaces. Let $J : U \to V$ be a Fréchet differentiable mapping. Assume the positive cone in V defined by $J(u) \geq 0$ contains an interior point. Let $F : U \to \mathbb{R}$ be a Fréchet-differentiable functional. Suppose u_0 satisfies*

$$F(u_0) = \min\{F(u) \mid u \in U,\ J(u) \geq 0\},$$

and that u_0 is a regular point of the inequality $J(u) \geq 0$. Then there exists $\eta \in V'$, such that

$${}_{U'}\langle F'(u_0), u\rangle_U = {}_{U'}\langle J'(u_0)^T \eta, u\rangle_U \text{ for all } u \in U. \tag{1.4}$$

Moreover,

$${}_V\langle J(u_0), \eta\rangle_{V'} = 0. \tag{1.5}$$

Equation (1.4) is the Euler-Lagrange equation, and is often an ordinary or partial differential equation, or an algebraic equation. In the particular case when U van V are function spaces, (1.5) is equivalent to the two identities

$$\eta \geq 0, \tag{1.6}$$

$$\operatorname{supp} \eta \subset \{x \in \mathbb{R} \mid {}_V\langle J(u_0(x)), \eta\rangle_{V'} = 0\}. \tag{1.7}$$

These two results are, compared to the Euler-Lagrange equation, equally important in the study of stationary points. They are very useful in proving characteristics of minimizers such as regularity and periodicity. One remarkable feature of Theorem 1.1 is that we have converted an inequality $F'(u_0) \cdot u \geq 0$ into an equality (1.4). This greatly facilitates the study of stationary points.

In elasticity theory, the main topic in this thesis, the Lagrange multiplier η may be interpreted as the 'force' necessary to enforce the constraint $J(u_0) \geq 0$. It only acts in *contact points*, i.e., points where $J(u_0(x)) = 0$, as indicated by (1.7).

Study of global minimizers: cut and paste

Among the family of stationary solutions there are generically few true minimizers. These global minimizers often have certain special properties such as additional symmetry that the other stationary solutions lack (the appearance of extra symmetry properties for global

minimizers is in fact one of the motivations behind the research in this thesis). Such properties may be studied by so-called 'cut and paste' arguments: we may modify any function, and study the change in energy as a result, as long as the modified function remains within the set of admissible functions. This is a very flexible and powerful way to isolate certain characteristics of global minimizers. In the two problems studied in this thesis in Chapters 2 and 5, we have to deal with the non-local nature of the constraints involved, and this makes cutting and pasting a challenge. Nonetheless, an application of this technique in a non-local setting has been achieved in Theorem 2.22.

Contact problems

A contact or obstacle problem is an energy minimization problem with a particular prescribed condition on the admissible functions. Typically, the functions have to satisfy an inequality rather than an equality. The classical example is that of a membrane that hits a table while hanging from a wire, illustrated in Figure 1.2. Let $u(x)$ denote the distance between the membrane and the table at position x. The energy of the membrane is given by

$$F(u) = \int_\Omega \left[|\nabla u|^2 + 2u \right]. \tag{1.8}$$

Here Ω is the domain of the membrane, defined by the wire located at the boundary $\partial\Omega$. The first term in (1.8) specifies contributions by elastic forces, the second takes gravity into account. The obstacle problem is then given by the following minimization problem:

$$\inf \left\{ \int_\Omega F(u) \;\middle|\; u \in H^1(\Omega),\ u \geq 0,\ u = g \text{ on } \partial\Omega \right\}. \tag{1.9}$$

The classical results on this problem indicate the main questions one needs to address to characterize the structure of the minimizers of general contact problems. We will briefly review them.

THEOREM 1.2 (Existence and regularity of the minimizer [48]). *There exists a unique minimizer u of problem (1.9). This minimizer satisfies*

$$\begin{cases} -\Delta u = 1 \text{ on } \{u > 0\} \subset \Omega, \\ u \in C^{1,1}(\bar{\Omega}). \end{cases}$$

The contact set $\omega_c := \{x \mid u(x) = 0\}$, has been exhaustively studied and has proved to have an intricate structure [18, 77, 19, 21, 141]. Here we merely cite two important results, on its area and on its convexity. The more interested reader will find a detailed exposition in [49].

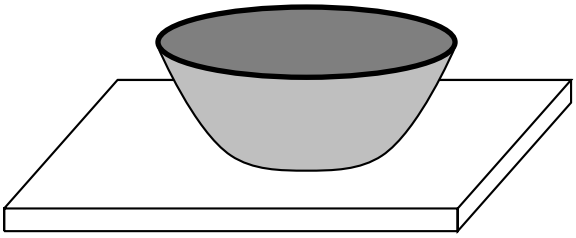


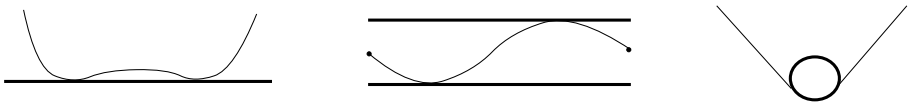
Figure 1.2: The classical obstacle problem: how does a membrane hang from a wire, and how does it interact with the table?

THEOREM 1.3 (Length of the boundary of the contact set [15, 20]). *There exists a constant C such that*

$$\mathcal{H}^1(\partial\omega_c) \leq C,$$

where $\mathcal{H}^1(V)$ is the 1-dimensional Hausdorff measure of a set V (see [42]*).*

THEOREM 1.4 (Convexity of the contact set [75]). *Let Ω be a smooth convex domain. For each $g > 0$, let u_g be a minimizer. Then the contact set $\omega_c(u_g)$ is convex and analytic.*



(a) (b) (c)

Figure 1.3: Three other examples of obstacle problems: (a) a piece of paper hitting a table while hanging down; (b) an elastic rod constrained by two parallel plates; (c) a membrane deformed by a heavy ball.

There are many other well-known examples of contact problems from mechanics. Three of these are illustrated in Figure 1.3: a piece of paper hitting a table as it hangs down, a wire that is held at the end points and is constrained by two parallel plates [71], or a membrane that is deformed by a rolling heavy ball it supports [9]. Basic questions that arise in each of these problem are:

(1) **How should we take contact into account to facilitate analysis?** In many cases (such as the classical obstacle problem) the model describing the physical phenomenon under scrutiny is well suited for analytical treatment. However, buckling rod problems—due to their complex geometrical shapes—form one class of problems where the contact constraint needs special attention. In Chapter 2

some work is done to obtain a contact condition in a simple form that allows further analysis.

(2) **When and how does contact develop?** As seen in Figure 1.4, the set of contact points often seems to have a particular simple structure. In the helically snarled rod, Fig. 1.4 (b), visual inspection suggests that contact is simply formed along one closed interval. Such simplicity does not always occur however. Two recent examples in which the contact set is counter-intuitive—despite the simple topology—are shown in Figure 1.4 (a) and (c)–(d). In (c) and (d), the problem under investigation is to find a *ropelength minimizing* configuration—rather than an energy minimizer—that takes (self-)contact into account with a prescribed topology. It has recently been shown that the contact set of the resulting minimizer consists of a single loop with four cusp points [24]. Intriguingly, the two components of the minimizing configuration do not touch at the centre point of symmetry: there is a gap of up to 6% of the diameter of the rods. In case of the single ply, shown in Fig. 1.4 (a), numerical results on closely related closed rods have shown that the contact set often consists of three components: apart from a single interval of contact points, there are two additional contact points near the loops at both ends [28, 53]. This suggests that for the single ply there often exists a single isolated contact point at the closing loop, next to the line of contact.

(3) **How does the contact constraint influence the regularity of the resulting solutions?** The classical obstacle problem (1.9) has $C^{1,1}$ minimizers (Theorem 1.2). The Gehring simple clasp in Figure 1.4 (c) is once more an example of surprising behaviour of the ropelength-minimizing configuration. The resulting solution is only $C^{1,2/3}$ (and is also in $W^{2,3-\varepsilon}$ for all $\varepsilon > 0$ but has no higher regularity), and the curvature of each tube blows up at the tip [24].

(4) **What are other general properties of minimizing solutions?** Examples here include the support of minimizers, their symmetry or their periodicity.

These four questions form the basis of our investigations in Chapters 2 through 5. We are now ready to derive the two minimization problems studied in those chapters. In Section 1.2 we model elastic rods on cylinders, and in Section 1.5 we introduce a model for capturing the formation of lipid bilayers. We also give a flavour of some of the problems that have to be overcome to obtain the desired results.

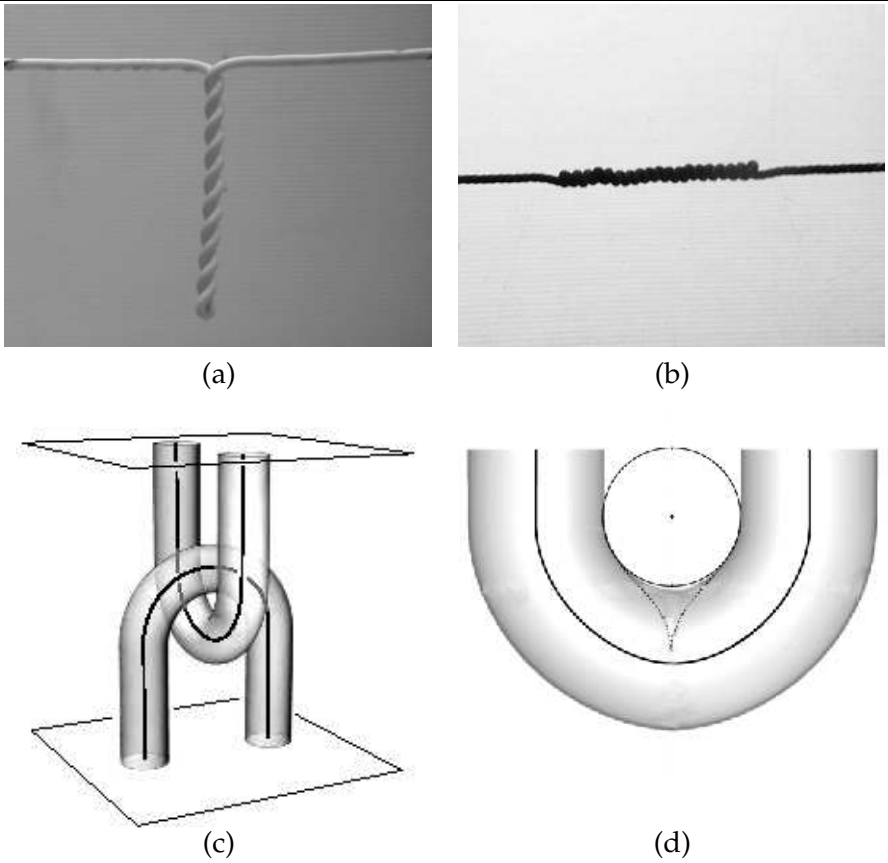


Figure 1.4: Three examples of elastic rods that exhibit different forms of self-contact. The two figures (a) and (b) exhibit rods buckling under different tensions: (a) is under low tension, (b) under high tension (a study of the bifurcation between these two states can be found in [64]). Visual inspection suggests that in both cases there is a single interval of contact. In the left rod, however, numerical evidence from circular rods seems to suggest the interval is supplemented with an isolated contact point at the looped end [27]. The bottom two figures (c) and (d) illustrate the Gehring simple clasp in perspective view ((c), not minimized for ropelength) and in orthogonal view ((d), minimized for ropelength) (adapted from [24]). In figure (d), the open round circle denotes the tube going through the shaded tube. Note the small but definite gap between the two tubes, centred at the origin of the two axes. The contact set, consisting of a loop with four cusps, is denoted by the dotted line. Only the front half can be seen in this picture.

1.2. Rods on cylinders

Elastic rods are idealizations of objects found in many physical situations, from shoe laces, to electricity wires, to DNA molecules. The manner in which they buckle can be both fascinating and problematic—depending on whether one is a scientist or whether one for instance tries to drill a hole to get to oil deep underground, and the drill string is buckling in all kinds of unwanted manners.

We will focus on a particularly simple type of elastic rod. Picturing a rod as a set of circular cross-sections along a centreline, we call a rod inextensible if the centreline does not extend or contract locally during deformation. It is termed unshearable if there are no discontinuous jumps in the orientation of the the cross-sections as we go along the centreline compared to the unbuckled state.

Now consider the following little experiment. Take an elastic rod—such as a piece of string, or a flexible electricity wire— by the ends and hold it between thumbs and fingers at each end. Extend the rod to its full length. Increase the internal twist by rotating the rod at one of the ends. As the hands are brought together, the rod starts to buckle, and one is likely to end up with a shape much like the single ply illustrated in Figure 1.4 (a).

We firmly believe that the single ply in Figure 1.4 (a) is the global minimizer for the minimization problem that corresponds to the above experiment, but we cannot prove this yet. To show why it is to be expected that this is potentially a hard problem, take again the simple ply. Now keep the two ends together in one hand without releasing the ply, and hold firmly. With the other hand, grab the loop at the bottom of the ply, and rotate it such that you unwind the ply. You will feel there is quite some tension within the rod. Now let it spring back and watch what happens. The result will be a configuration such as in Figure 1.5. This is most probably not a global minimizer, but only a local one. However, it is very stable, in the sense that violently shaking the rod does not change its shape. Mathematically we will have to prove that we can improve on such local minimizers, and this is no mean task. The cause of the multitude of local minimizers is that the rod cannot physically intersect itself. This means the state space of possible rod configurations has a very complicated structure, making cut and paste arguments very difficult.

As a first step towards mathematically proving theorems about rods that are free to buckle in any direction, we will consider a rod

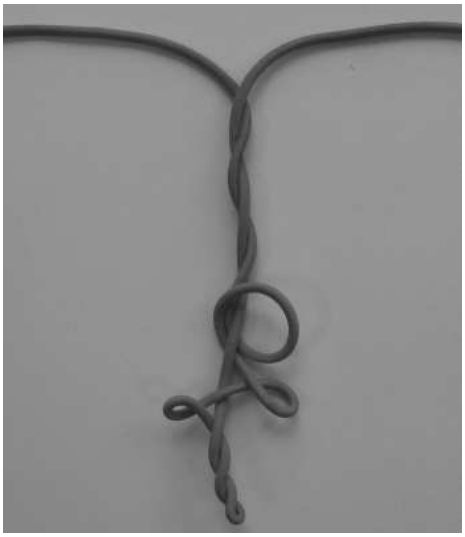


Figure 1.5: A rod configuration that is not a global minimizer, but only a local minimizer. Such configurations are quite stable under perturbation and geometrically diverse, indicating that minimizing elastic energy among all free rods poses a significant challenge.

which is constrained to lie on a cylinder. This problem has a lower dimensionality, making it more accessible for analysis.

Dead and rigid loading

Engineers distinguish two ways in which a body may be loaded by external forces: through dead and through rigid loading.

In a dead loading experiment, the *amount of force* is specified, and has some displacement associated with the force as a result. For example, we may push on an elastic spring with a force of 1 N, with a shortening of the spring of 1 cm as a result.

In rigid loading, on the other hand, it is the *displacement* which is predetermined in the experiment. The loading force is then an dependent quantity. The elastic rod is a good example here: with the rod between our thumbs and fingers, it is easy to twist the ends relative to each other by half a turn. The applied force is now unknown, but could be measured in an experimental setup.[1]

[1] The German terms for dead and rigid loading, *Laststeuerung* (steering through force) and *Wegsteuerung* (steering through displacement), indicate the difference between the two terms particularly well.

The important difference between the two loading conditions is that their stationary points are identical, but their stability (in a variational sense) may differ.

Modeling rods on cylinders

Consider an inextensible, unshearable, elastic rod of circular cross-section that is constrained to lie on a cylinder, and which is subject to a force T and a moment M at the ends (see Figure 1.6). We assume that at the rod ends, T, and M are maintained parallel to the axis of the cylinder, but that the loading device leaves the rod ends free to rotate around the circumference of the cylinder; the ends of the rod therefore need not be coaxial. The rod is naturally straight and inextensible, and cross-sections are assumed to remain orthogonal to the centreline. We will derive a minimization problem for rods of length 2ℓ and later take the limit $\ell \to \infty$.

In the Cosserat rod theory [7, Ch. VIII] the configuration of this rod is characterized by a right-handed orthogonal rod-centred coordinate frame of directors, $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3\}$, each a function of the arc length parameter s (see Fig. 1.6). The director $\boldsymbol{d}_3$ is assumed parallel to the centreline tangent, and by the assumption of in extensibility the centreline curve $\boldsymbol{r}$ satisfies

$$\dot{\boldsymbol{r}} = \boldsymbol{d}_3,$$

where the dot denotes differentiation with respect to arclength. The strain of the rod is characterized by the vector-valued function $\boldsymbol{u}$ given by

$$\dot{\boldsymbol{d}}_k = \boldsymbol{u} \times \boldsymbol{d}_k, \qquad k = 1, 2, 3.$$

When decomposed as $\boldsymbol{u} = \kappa_1 \boldsymbol{d}_1 + \kappa_2 \boldsymbol{d}_2 + \tau \boldsymbol{d}_3$, the components may be recognized as the two curvatures $\kappa_{1,2}$ and the torsion τ.

We choose a fixed frame of reference $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$, where $\boldsymbol{e}_3$ is parallel to the cylinder axis, and we relate the frame $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3\}$ to this frame by a particular choice of Euler angles $\{\theta, \psi, \phi\}$ [69, 65]. In this parametrization θ is the angle between $\boldsymbol{d}_3$ and $\boldsymbol{e}_3$ (or between the centreline and the cylinder axis), ψ characterizes the rotation around the cylinder axis, and ϕ is a partial measure of the rotation between cross-sections. The condition that the centreline of the rod lie on the surface of a cylinder of radius r translates into the kinematic condition

$$\dot{\psi} = \frac{1}{r} \sin\theta.$$

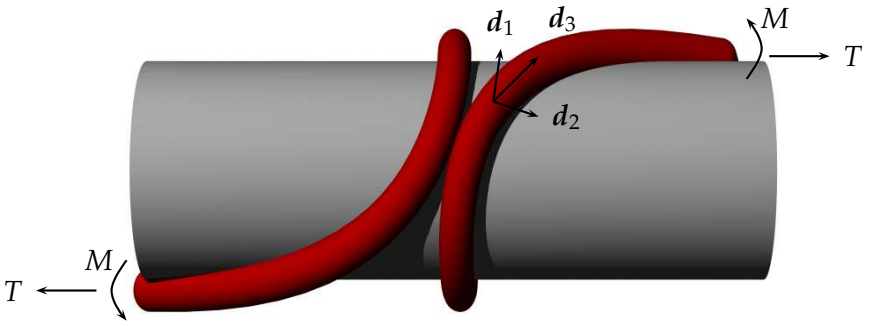


Figure 1.6: Modeling the rod on a cylinder. A tension T and moment M is applied at the ends. The three arrows attached to a point on the rod designate the frame of directors $\{d_1, d_2, d_3\}$.

In terms of the remaining degrees of freedom $\{\theta, \phi\}$ the curvatures and torsion are given by

$$\begin{aligned}\kappa_1 &= \dot{\theta}\sin\phi - \frac{1}{r}\sin^2\theta\cos\phi,\\ \kappa_2 &= \dot{\theta}\cos\phi + \frac{1}{r}\sin^2\theta\sin\phi,\\ \tau &= \dot{\phi} + \frac{1}{r}\sin\theta\cos\theta.\end{aligned}$$

The variational problem

For a given rod the stored elastic energy is a sum of curvature and torsional energy, and is given by [69],

$$\begin{aligned}E(\theta,\tau) &= \frac{B}{2}\int_{-\ell}^{\ell}(\kappa_1^2+\kappa_2^2) + \frac{C}{2}\int_{-\ell}^{\ell}\tau^2\\ &= \frac{B}{2}\int_{-\ell}^{\ell}\dot{\theta}^2 + \frac{B}{2r^2}\int_{-\ell}^{\ell}\sin^4\theta + \frac{C}{2}\int_{-\ell}^{\ell}\tau^2. \qquad (1.10)\end{aligned}$$

Here B is the bending stiffness of the rod, C the torsional stiffness, and r the radius of the cylinder. In this entire section, all integrals refer to integrals with respect to arclength. To determine the work done by the tension and moment at the ends of the rod we need to characterize the displacements associated with these forces. For the tension T the associated displacement is the shortening S,

$$S(\theta) = \int_{-\ell}^{\ell}(1-\cos\theta).$$

The generalized displacement associated with the moment M is the end rotation R, which is well-defined by the assumption of constant end tangents. We assume rigid loading in shortening and dead loading in torsion, i.e., we prescribe the shortening S and the moment M, which implies that the tension T and the end rotation R are unknown and to be determined as part of the solution.

This loading condition leads to the minimization problem

$$\inf\{E(\theta,\tau) - MR(\theta,\tau) \mid S(\theta) = s\}. \tag{1.11}$$

It is quite common to replace the end rotation R by a topologically inspired quantity called *link*. This concept is traditionally defined only for *closed rods*—rods whose end points meet—and as such not accessible for use in the present context.

We will spend the entire Chapter 3 making this transition mathematically precise, and therefore postpone a proper introduction into this subject to Section 1.4.

1.3. The contact problem

Let us for now assume we have such a a precise identification between end rotation R and a quantity termed 'open link' Lk^o. We are then in a position to properly study the contact problem for rods on cylinders. Recall that we wish to minimize (1.11).

In Euler angles, the open link is given by

$$Lk^o = \int_{-\ell}^{\ell} (\dot{\phi} + \dot{\psi}) = [\phi + \psi]_{-\ell}^{\ell}.$$

In Chapter 3 we prove that the identification between end rotation and open link is correct in an open set around the undeformed configuration $\theta \equiv 0$, but loses validity when $|\theta|$ takes values larger than π. Although nothing we have seen suggests that in an energy-minimizing situation θ would take values outside of the admissible interval $(-\pi, \pi)$, we have no rigorous argument to guarantee that θ remains inside that interval, and therefore we are forced to assume this. In terms of the variables θ and τ this functional then takes the form

$$Lk^o(\theta,\tau) = \int_{-\ell}^{\ell} \left(\tau + \frac{1}{r} \sin\theta(1 - \cos\theta) \right).$$

We may now subsitute (1.11) for

$$\min\{E(\theta,\tau) - MLk^o(\theta,\tau) : S(\theta) = s\}$$

for given $s > 0$. The tension T has a natural interpretation as a Lagrange multiplier associated with the constraint of S.

We can simplify this minimization problem by first minimizing with respect to τ for fixed θ, from which we find $\tau \equiv M/C$; re-insertion yields the final minimization problem

$$\min \{F(\theta) : S(\theta) = s\} \tag{1.12}$$

with

$$F(\theta) = E(\theta) - MLk^o(\theta) = \\ \frac{B}{2}\int \dot{\theta}^2 + \frac{B}{2r^2}\int \sin^4\theta - \frac{M}{r}\int \sin\theta(1-\cos\theta). \tag{1.13}$$

We are interested in localized forms of deformation, in which the deformation is concentrated on a small part of the rod and in which boundary effects are avoided, and therefore we take an infinitely long rod and consider θ, F, and S to be defined on the whole real line. From the construction above it is then natural to assume that $\theta \to 0$ as $|s| \to \infty$, and this limit behaviour is consistent with the form of F.

Behaviour of minimizers

The Euler-Lagrange equations associated with the minimization problem (1.12) can be written as a Hamiltonian system with one degree of freedom,

$$\frac{1}{2}\dot{\theta}^2 + V(\theta) = H. \tag{1.14}$$

In this system two independent parameters remain, which may be interpreted as a scaled cylinder radius $\tilde{r} = rM/B$ and a combined loading parameter $m = M/\sqrt{BT}$.

Solutions of the original minimization problem are orbits of this Hamiltonian system that are homoclinic to zero, and such orbits have been studied in detail in [65]. Among the findings are:

(1) for all values of $\tilde{r}$ ranges of m exist with orbits that are homoclinic to the origin;
(2) at some parameter points these homoclinic orbits 'collide' with saddle points. The saddle points correspond to helical solutions (constant angle θ) and close to these collisions the homoclinic orbit has a large region of near-constant angle θ.

In Figure 1.7 a bifurcation diagram is shown with two such collisions, one at a forward helix ($0 < \theta < \pi/2$, at $m = m_{c_2}$) and one at a backward helix ($\pi/2 < \theta < \pi$).

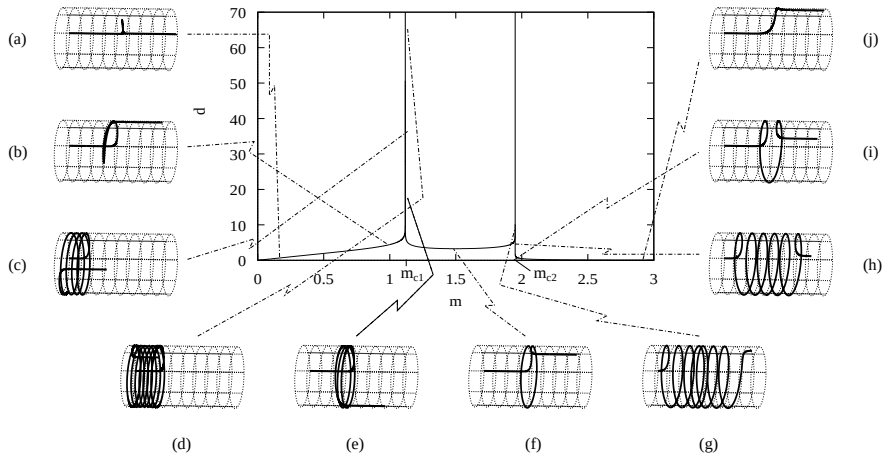


Figure 1.7: A load-displacement diagram showing shortening d of stationary points as a function of the (combined) load m (from [65]). Contact effects are not taken into account. The peaks divide this diagram into three sections. The solutions in the middle section intersect themselves, whilst the solutions on the right do not. The section on the left consists of heteroclinic connections between $\theta = 0$ and $\theta = 2\pi$ which are not considered here. For sufficiently large shortening, the rod configuration that has lowest energy is on the self-penetrating branch.

In [65] the question of stability of these solutions, both local and global, was left untouched. The framework of [92] suggests that in each 'peak' the right-most curve is locally stable; in the case of two peaks this does not give any argument to distinguish between the two.

In Chapter 2 we focus on global energy minimization. We start the discussion by showing that for sufficiently large shortening, and when neglecting contact effects, *global energy minimizers always intersect themselves*.

It is this result that gives rise to our study: since energy minimization without an appropriate non-self-intersection condition leads to self-intersection, such a condition is necessary to study physically correct solutions.

The main minimization we study is

Find a function $\bar{u} \in J$ such that

$$F(\bar{u}) = \min\{F(u) \mid u \in J\}, \tag{1.15}$$

where

$$F(u) = \int_0^T a(u(x))u'^2(x) + b(u(x))dx,$$

and

$$J = \Big\{ u \in 1 + H_0^1(0,T),\ \int_x^{x+1} u \geq 0\ \forall x \in [0, T-1] \Big\}.$$

The functions a and b are taken as general as possible.

The main results in Chapter 2 are the derivation of the Euler-Lagrange equations of (1.15), and the full characterization of their solutions. The results culminate in an expression of great simplicity of the contact set:

THEOREM 1.5. *Let u be a minimizer of (1.15), with corresponding contact set*

$$\omega_c := \Big\{ x \in [0, T-1] \ \Big|\ \int_x^{x+1} u = 0 \Big\}.$$

Then ω_c is an interval.

1.4. Link, twist, writhe, and end rotation

A large part of this thesis considers the buckling of rods under forces such as end tension and end moments. The study of these structures requires a set of concepts with which we can distinguish different rod configurations. In particular, one would like to have measures indicating how contorted or buckled a given rod is, both in a quantitative and in a qualitative way. The strain energy (1.10) is of course one of these: the higher this energy, the more contorted the rod. Historically, geometric and topological measures have been used to quantitatively describe rod configurations. We will review the three most commonly used, and their relation to the strain energy already mentioned. Two of these measures are defined for *closed rods* only. As our basic objects of study are *open rods*, it would enhance our insight into the buckling behaviour of these rods if we had analogous concepts for this class. The development of such concepts, and the study of their applicability is the subject of Chapter 3.

For our purposes a rod is a member of the set

$$\mathcal{A}^0 = \{\ (\boldsymbol{r}, \boldsymbol{d}_1) \in C^2([0,L]; \mathbb{R}^3 \times S^2) \text{ such that } |\dot{\boldsymbol{r}}| \neq 0,\ \dot{\boldsymbol{r}} \cdot \boldsymbol{d}_1 = 0, \text{ and } \boldsymbol{r} \text{ is non-self-intersecting } \}.$$

A closed rod is an element of $\mathcal{A}^0$ for which the beginning and end connect smoothly.

Recall that at each point on the centerline of the rod we attach a right-handed orthonormal frame $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3\}$. The twist of a closed rod $(\boldsymbol{r}, \boldsymbol{d}_1)$ is now defined by

$$Tw(\boldsymbol{r}, \boldsymbol{d}_1) := \frac{1}{2\pi} \oint_{\boldsymbol{r}} \dot{\boldsymbol{d}}_1(s) \cdot \boldsymbol{d}_2(s)\, ds. \tag{1.16}$$

It measures the number of times $\boldsymbol{d}_1$ revolves around $\boldsymbol{d}_3$ in the direction of $\boldsymbol{d}_2$ as we go around the rod.

Let $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ be two non-intersecting closed curves. Then the link of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ is defined by

$$Lk(\boldsymbol{r}_1, \boldsymbol{r}_2) := \frac{1}{4\pi} \oint_{\boldsymbol{r}_1} \oint_{\boldsymbol{r}_2} \frac{[\dot{\boldsymbol{r}}_1(s) \times \dot{\boldsymbol{r}}_2(t)] \cdot [\boldsymbol{r}_1(s) - \boldsymbol{r}_2(t)]}{|\boldsymbol{r}_1(s) - \boldsymbol{r}_2(t)|^3}\, dsdt. \tag{1.17}$$

The writhe of a closed curve r is

$$Wr(\boldsymbol{r}) := \frac{1}{4\pi} \oint_{\boldsymbol{r}} \oint_{\boldsymbol{r}} \frac{[\dot{\boldsymbol{r}}(s) \times \dot{\boldsymbol{r}}(t)] \cdot [\boldsymbol{r}(s) - \boldsymbol{r}(t)]}{|\boldsymbol{r}(s) - \boldsymbol{r}(t)|^3}\, dsdt. \tag{1.18}$$

In Figure 1.8 we give some typical numbers for link and writhe for specific closed rods.

The link, twist and writhe of a closed rod are related by the well-known Călugăreanu-White-Fuller Theorem [22, 142, 50]:

THEOREM 1.6. *Let $(\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^0$ be a closed rod as defined above. Then*

$$Lk(\boldsymbol{r}, \boldsymbol{d}_1) = Tw(\boldsymbol{r}, \boldsymbol{d}_1) + Wr(\boldsymbol{r}). \tag{1.19}$$

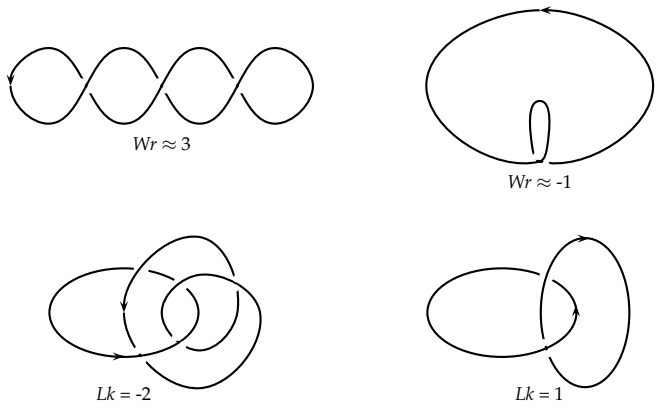


Figure 1.8: Examples of linking and writhing numbers for different (pairs of) curves.

An illustration of Theorem 1.6 is shown in Figure 1.9. Since link is a topological invariant, twist and writhe are converted into one another during continuous deformations of an elastic rod. One example where this theorem has been useful is the well-known supercoiling behaviour of DNA. Enzymes manipulate the link of the DNA, by cutting and gluing it, while inserting extra twist. To find a new energy minimum, the DNA starts to coil, which decreases twist, but increases writhe. This phenomenon forms the basis of the tight packing of DNA into neat chromosomes.

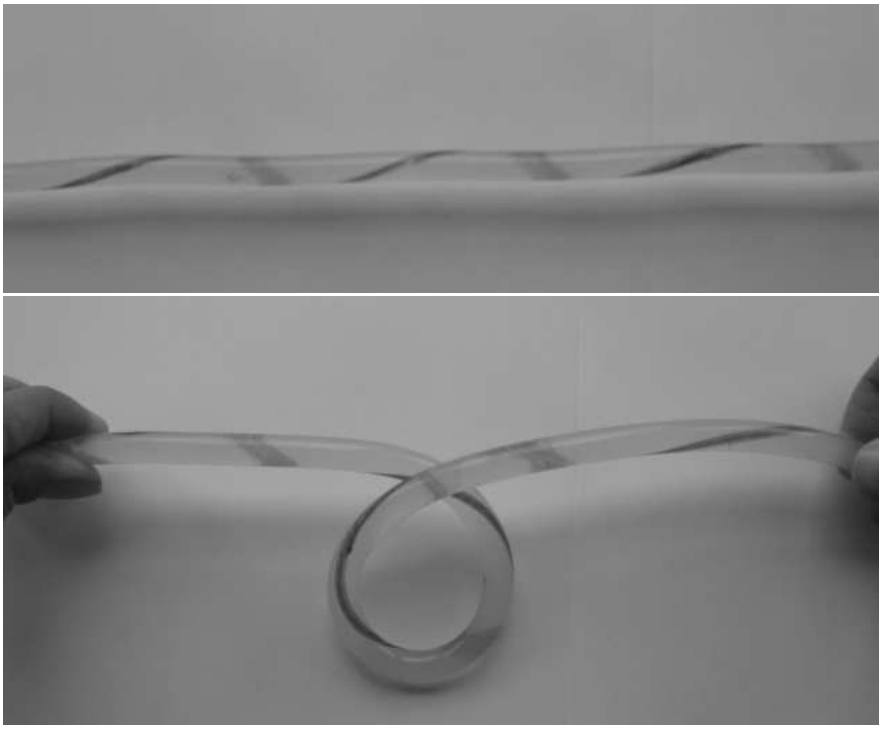


Figure 1.9: Illustration of $Lk = Tw + Wr$. Starting with a rod configuration with high twist (top), the rod is deformed keeping link constant (i.e., keeping the ends clamped between thumb and fingers). The resulting rod (bottom) has lower twist, but higher writhe.

In Chapter 3 we define new functionals Lk^o, Tw^o, and Wr^o for open rods for a specific class of functions $\mathcal{A}^1_{r_0}$. The main result states that these functionals are well-defined for open rods in $\mathcal{A}^1_{r_0}$. An immediate

and philosophically pleasing corollary of these definitions is that we recover Theorem 1.6 for open rods in $\mathcal{A}^1_{r_0}$:

THEOREM 1.7. *Let $(\boldsymbol{r}, \boldsymbol{d}_1)$ be a rod in $\mathcal{A}^1_{r_0}$. Then*

$$Lk^o(\boldsymbol{r}, \boldsymbol{d}_1) = Tw^o(\boldsymbol{r}, \boldsymbol{d}_1) + Wr^o(\boldsymbol{r}).$$

The initial objective for the energy minimization problem of rods on cylinders—to show that end rotation can indeed be equated with an open link functional—is also proved, for a slightly different class of functions, $\mathcal{A}^2_{r_0}$:

THEOREM 1.8. *Let $(\boldsymbol{r}, \boldsymbol{d}_1)$ be an open rod in $\mathcal{A}^2_{r_0}$. Then*

$$R(\boldsymbol{r}, \boldsymbol{d}_1) = 2\pi Lk^o(\boldsymbol{r}, \boldsymbol{d}_1). \tag{1.20}$$

We end Chapter 3 with a detailed discussion on the particular definition of the class of functions $\mathcal{A}^1_{r_0}$. We show it cannot be enlarged by simply doing away with one of its nine defining conditions: we give appropriate examples that show that Wr^o becomes ill-defined in each of these cases.

1.5. Lipid membranes

Cells are the main building blocks of life. They are small factories filled to the brim with a wonderful array of machinery, such as proteins, genetic material, mitochondria, and so on. For a cell to function, its contents need to be separated from that of other cells. This task is performed by lipid bilayers. They can in fact be found in many places other than the cell's main perimeter.

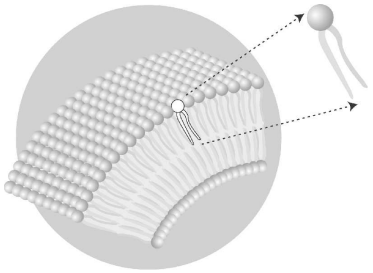


Figure 1.10: A membrane is a bilayer of lipid molecules.

Lipid membranes have some unique properties. Their main ingredient is a lipid molecule with a head and two tails. The absence

of covalent bonds between the lipid molecules is responsible for the most salient property of lipid membranes: their flexibility and elasticity. When pressed, they bounce back upon release. This also enables the occurrence of vesicles, small compartments pinched off of a larger membrane structure that are used for transport of substances inside the cell.

Current modeling based on reaction-diffusion does not yet allow us to study the elastic properties of lipid membranes. Here we study a very simplified model, which is based on the framework of Fraaije and coworkers [45], that will eventually allow for the investigation of elastic properties of lipid bilayers. This model has been fully derived in [11].

The main motivation behind this research is the following phenomenon. When adding a sufficient quantity of lipid molecules into a container with water, the lipids start to spontaneously aggregate into clusters. Such clusters are called micelles and consist of a single layer of lipid molecules with the heads pointing outward. Micelles start to form above a particular concentration, called the the Critical Micelle Concentration (CMC, see e.g., [54, Chapter 1]). This is illustrated in Figure 1.11.

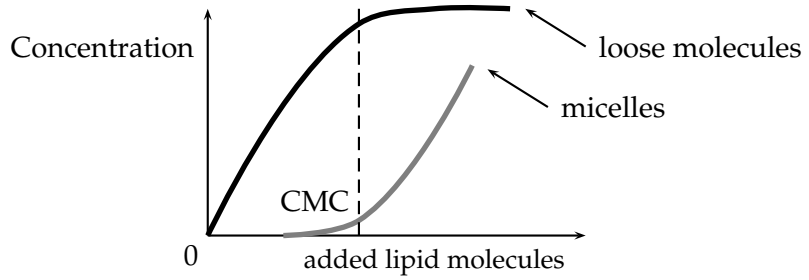


Figure 1.11: The Critical Micelle Concentration: while adding more lipid mass to a container filled with water, micelle membranes start to form spontaneously for lipid concentration above the CMC.

We will try to capture this difference in spreading or aggregation using energy minimization techniques, in order to show that models based on a continuum description have a good reason to be studied in this context.

Derivation of a model for lipid bilayers

The main ingredient of lipid membranes is the lipid molecule. It is made from a head and two tails. The head is hydrophilic, and readily dissolves in water. The tails, however, are hydrophobic, i.e., there exist

repelling forces between the tails and water. The lipid molecules are modelled as two beads with a stiff rod of fixed length between them (see Figure 1.12). Membranes are formed of two layers of these lipids that are arranged with the tails towards each other (see Figure 1.10). They are largely surrounded by water molecules, which we also model by single beads.

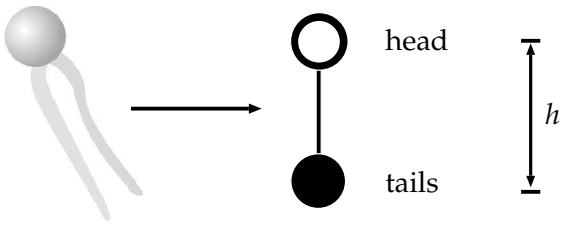


Figure 1.12: The model for a lipid molecule: a head and tail bead with a stiff rod of fixed length h between them.

We consider a one-dimensional representation of the membrane, where the single spatial coordinate is perpendicular to the plane of the membrane. We regard the lipid molecules as essentially rigidly oriented, either pointing with their head to the positive or negative side of the axis, and hence we identify two distinct species of lipid molecules.

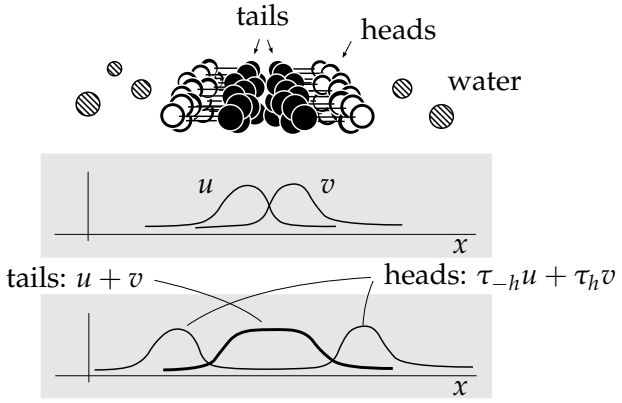


Figure 1.13: Example distributions of left-oriented tails u and right-oriented tails v, and the corresponding distributions of the respective heads, $\tau_{-h}u$ and τv.

Their densities are given by u, and v, or more specifically, u and v denote the densities of the *tail beads*. If we model a lipid molecule as

a head and tail bead that are a distance h apart, then the densities of head beads of the two kinds of molecules is given by $\tau_{-h}u$ and $\tau_h v$ respectively. This is illustrated in Figure 1.13. Lastly, let w be the density of water molecules. We assume all beads are of equal size, so we can speak of volume fractions.

The different molecules influence each other, both through entropic effects and through hydrophobic or hydrophilic interactions. We hypothesize that the combination between entropy, which causes spreading of mass, and interaction energy, which causes congregation of lipid mass, gives rise to membrane-like structures. To investigate this hypothesis, we define the following energy functional

$$T\int [s(u)+s(v)+s(w)]+\alpha\int (w+\tau_{-h}u+\tau_h v)\,\kappa*(u+v), \qquad (1.21)$$

where

$$s(t)=\begin{cases} t\log t & t\geq 0,\\ 0 & t<0.\end{cases}$$

The first integral in (1.21) denotes the Gibbs-Boltzmann entropy, multiplied by the temperature T. The second integral represents the interaction between the different beads, and penalizes proximity between the tails on the one hand and the heads and water on the other. The interaction kernel κ chosen here is

$$\kappa(t)=\frac{1}{2}e^{-|t|}.$$

This choice is based on laboratory experiments [72, 139], and allows for much analysis.

We assume that the solution is incompressible. The total fraction of beads is scaled to 1, giving rise to the following relation

$$w=1-u-v-\tau_{-h}u-\tau_h v.$$

Since w is non-negative, we have the following contact condition

$$u+v+\tau_{-h}u+\tau_h v\leq 1. \qquad (1.22)$$

The free energy in (1.21) is now reduced to

$$\int [s(u)+s(v)]+\alpha\int (1-u-v)\,\kappa*(u+v). \qquad (1.23)$$

Here the temperature T is scaled into α. Note that we have discarded the entropy contribution of the water molecules. We assume this simplification has little influence on the results.

The formal mathematical problem that follows from this discussion, is the following minimization problem.

Find functions $(\bar{u}, \bar{v}) \in C$ such that

$$G(\bar{u}, \bar{v}) = \min\{F(u,v) \,|(u,v) \in C\}, \tag{1.24}$$

where

$$G(u,v) = \int [s(u) + s(v)] + \alpha \int [(1-u-v)\,\kappa * (u+v) - \zeta(c_0)],$$

and

$$C = \{(u,v) \in (c_0, c_0) + L^1(\mathbb{R})^2 \mid u + v + \tau_{-h}u + \tau_h v \leq 1 \text{ a.e.}\}.$$

Here, $\zeta(c_0) = 2s(c_0) + \alpha c_0(1-2c_0)$ is added to ensure integrability.

Blom and Peletier [11] have recently studied this minimization problem, and have shown that minimizers indeed show a membrane-like structure for certain parameter values. One of the crucial results lacking from their analysis was to prove existence of minimizers for (1.24).

As a first step towards proving existence for (1.24), we study the existence of minimizers for the following simplified 1-dimensional minimization problem in Chapters 4 and 5:

Find a function $\bar{u} \in K_m$ such that

$$H(\bar{u}) = \min\{H(u) \mid u \in K_m\}, \tag{1.25}$$

where

$$H(u) = \int_{\mathbb{R}} \big[(u + c_0)\log(u/c_0 + 1) - u - \alpha u\,\kappa * u\big],$$

$$K_m = \Big\{u \in L^1(\mathbb{R}) \;\Big|\; \int_{\mathbb{R}} u = m,\ u \geq 0,\ u + \tau_h u \leq 1\Big\},$$

$$\kappa(x) = \frac{1}{2}e^{-|x|}$$

$$c_0 \in \Big(0, \frac{1}{2}\Big),\ \alpha > 0,\ h > 0.$$

To give a flavour of some of the major obstacles that have to be overcome, we recall the Concentration-Compactness Lemma.

LEMMA 1.9 (Lions, [86]). *Let $\{u_n\}$ be a sequence of positive functions in $L^1(\mathbb{R})$, with fixed mass $m > 0$. Then there exists a subsequence u_{n_k} satisfying one of three possibilities:*

(1) **vanishing**: *for all $L < \infty$,*

$$\lim_{k\to\infty} \sup_{y\in\mathbb{R}} \int_{y-L}^{y+L} u_{n_k} = 0;$$

(2) **dichotomy**: *there exists* $\mu \in (0, m)$ *such that for all* $\varepsilon > 0$ *there exist* $k_0 \geq 1$ *and* $\{v_k^1\}, \{v_k^2\} \subset L^1(\mathbb{R})$ *satisfying for all* $k \geq k_0$,

$$\begin{cases} \left\| u_{n_k} - (v_k^1 + v_k^2) \right\|_{L^1} \leq \varepsilon, \\ \left| \int v_k^1 dx - \mu \right| \leq \varepsilon, \\ \left| \int v_k^2 dx - (m - \mu) \right| \leq \varepsilon, \\ d(\operatorname{supp} v_k^1, \operatorname{supp} v_k^2) \to \infty \text{ as } k \to \infty; \end{cases}$$

(3) **compactness**: *there exists* $\{y_k\} \subset \mathbb{R}$ *such that*

$$\forall \varepsilon > 0 \; \exists L < \infty \text{ such that } \int_{y_k - L}^{y_k + L} u_{n_k} \geq m - \varepsilon.$$

The different possibilities are illustrated in Figure 1.14. The first two options in Lemma 1.9 are problematic when taking limits: mass 'escapes' to infinity, and the limit of the (sub)sequence does not lie in the original set of functions. To prove the existence of a minimizer of (1.25), we take a minimizing sequence in K_m. The goal is to show that there exists a subsequence that is compact in the sense of Lemma 1.9, since this assures the existence of a minimizer [86]. Often one can only prove this by showing that *all* minimizing sequences are compact, through exclusions of the two other options. There are two features in (1.25) that make this non-trivial:

- The non-local nature of the $\int u \, \kappa * u$ term in $H(u)$: this makes it difficult to compare energies of solutions that differ by a cut and paste operation.
- The non-local contact constraint $u + \tau_h u \leq 1$: for any new function we create from an admissible function using cut and paste operations, we have to make sure it is still admissible. In particular, it should still satisfy the contact constraint.

(a) (b) (c)

Figure 1.14: The three possibilities in the Concentration-Compactness Lemma 1.9: (a) vanishing, (b) dichotomy, and (c) compactness. The bold line in (c) illustrates the limit function. Arrows indicate taking limits.

Critical Micelle Concentration and unbounded domains

The preceding discussion raises the question why we study this minimization problem taking functions on an unbounded domain in the first place. After all, since the functional $H(u)$ is lower semicontinuous in an appropriate topology, existence of minimizers is guaranteed if we would take a compact domain!

One of the main features we want to address with the model in (1.25) is the occurrence of the Critical Micelle Concentration discussed before. How can we capture the division between either aggregation of lipid molecules into micelles, and evenly spreading out of these lipids? (Note that in the current one-dimensional model there is no distinction between a single-layered micelle and a double-layered membrane: both are made of lipid molecules whose tails point towards each other. This is of course not the case in higher dimensions.)

The very possibility that minimizers might *not* exist allows one to study the dichotomy between spreading and aggregation in the physical system. Non-existence correlates precisely with the situation in which the lipid molecules spread because there isn't sufficient lipid mass to spontaneously form micelles.

Results

The study of existence of minimizers is greatly dependent on having sufficient control over the admissible functions. Chapter 4 is therefore entirely devoted to the study of regularity of minimizing solutions.

In minimization problems that are not constrained by inequalities regularity of minimizers often follows directly from the Euler-Lagrange equations. For instance, if we consider

$$\inf\left\{H(u) \mid u \in L^1(\mathbb{R}),\ \int u = m\right\},$$

then a minimizer u satisfies

$$\log(u/c_0 + 1) - 2\alpha\,\kappa * u = \lambda,$$

for some $\lambda \in \mathbb{R}$. If we now assume that $u \in W^{k,p}$ then we can immediately conclude that $u \in C^\infty$ by a bootstrap argument: $\kappa * u \in W^{k+2,p}$, and therefore $u \in W^{k+2,p}$, and so on. We are, however, confronted with two inequality constraints, and hence additional Lagrange multipliers. Moreover, we will see from the derivation of the Euler-Lagrange

equations that these new multipliers have lower regularity than the admissible functions, and hence no knew knowledge on the smoothness of minimizers is gained.

We hence set out to study the regularity properties of minimizing solutions without using the Euler-Lagrange equations. The main result is the following theorem.

THEOREM 1.10. *Let u be a minimizer of (1.25). Then $u \in H^1(\mathbb{R})$. In particular, u is continuous.*

The continuity of minimizers allows for much more control than mere L^1-regularity, and forms a major ingredient in the study of the existence of minimizers, which is the subject of Chapter 5.

The next step is the derivation of the Euler-Lagrange equations. The method of proof is similar to that in Chapter 2, but additional care has to be taken to extend the proof to an unbounded domain and to incorporate multiple constraints. The H^1-regularity and the Euler-Lagrange equations together form the two cornerstones of the existence theorems.

We also briefly digress to study a simplified minimization problem, in which the upper contact condition $u + \tau_h u \leq 1$ has been dropped. This allows for symmetrization techniques which facilitate the study of minimizers. Existence for this simpler problem also forms the start of a proof that shows that minimizers exist for problem (1.25).

As discussed in great detail in Chapter 5, existence of minimizers for some particular $m > 0$ is assured if the function $I(m)$ defined by

$$I(m) := \inf\{H(u) \mid u \in K_m\}$$

satisfies

$$I(m) < I(\mu) + I(m - \mu) \text{ for all } \mu \in (0, m). \tag{1.26}$$

To show when this equation holds, much attention is given to the study of $I(m)$. We show that, depending on parameter choices, there are two qualitatively different cases, which are illustrated in Figure 1.15.

The main theorem of Chapter 5 is the following.

THEOREM 1.11. *Suppose that $2\alpha c_0 > 1$. Then for any $m > 0$ there exists a minimizer for (1.25).*

The above theorem holds only for specific values of α and c_0. For general parameter values, we show that we should not expect minimizers to exist for all $m > 0$. For values of α and c_0 such that $2\alpha c_0 < 1$, the existence of minimizers for (1.25) is investigated numerically. These

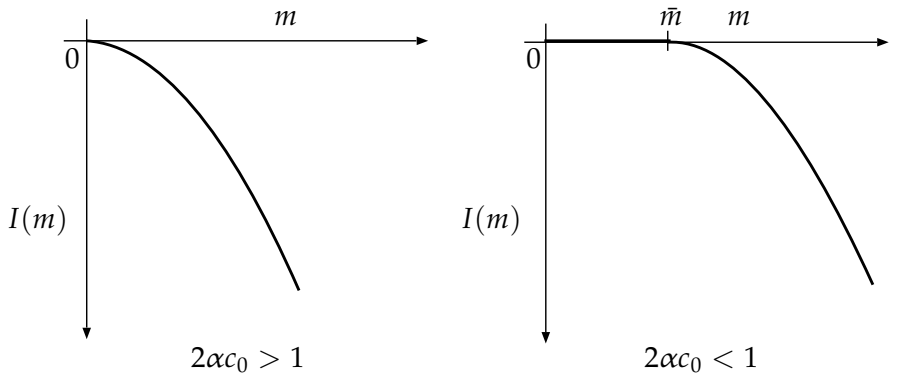


Figure 1.15: Two different cases of existence. If $2\alpha c_0 > 1$ (left), $I(m) < 0$ for all $m > 0$, and (1.26) holds. Hence minimizers exist for all $m > 0$. If $2\alpha c_0 < 1$ (right), $I(m) = 0$ for $m \in [0, \bar{m}]$ ($\bar{m} > 0$, and may depend on the other parameters), and $I(m) < 0$ for all $m > \bar{m}$. This suggests that minimizers do not exist at all for $m \in (0, \bar{m}]$.

computations also provide some insight into the general structure of minimizers.

1.6. Defence strategies against cuckoo parasitism

The last chapter is devoted to the study of a parasite-host system, and in particular focuses on the ways hosts may defend themselves against a parasite. We will take both evolutionary and ecological aspects into account. To set the stage, let us first consider some of the conditions under which evolution operates.

Constraints on evolution

The process of evolution by natural selection takes place within a set of constraints, in much the same way as for instance buckling in rods is constrained by a kinematic or contact condition. Such restrictions may be loosely divided in three categories.

- **Nutritional and physical constraints.** Shortages in food sources such as oxygen, water, temperature, or others may severely restrict growth and development of organisms, and constrain the evolutionary process by limiting the variety of organisms on which natural selection acts.
- **Genetic material.** All living organisms possess genetic material which is transfered from one generation to the next. Evolution

is therefore an inherently historical process in which the available DNA sets the stage for future adaptations. Most notably, the cascades of genetic networks that are expressed during the development of an organism cannot change drastically in short time spans. This poses important limitations on organisms regarding their ability to adapt to changing environments.

- **Trade-off laws.** In many instances one may find that different species have complementary traits, that are the outcome of a trade-off between their costs and benefits. Some well-known examples include
  (1) Nutritional value (size) of plant seeds vs. their capacity to disperse. Plant seeds are often either light and therefore have great dispersional abilities, or they are heavy, allowing them to have a better survival potential [137]. Another example is the growth rate of trees vs. the quantity of seeds they produce. Here energy constraints induce a trade-off between how much a plants invests in its own growth and how much it invests in the production of potential offspring [41].
  (2) Virulence of parasites. Parasites never benefit from the death of their hosts, and therefore virulence levels are often found to be relatively moderate [5, 96]. Current theory predicts the existence of an ongoing trade-off between transmission of the parasite on the one hand and virulence (host mortality) or clearance (host immune reaction to the parasite) on the other. This results in the optimal fitness of the parasite given its life history characteristics. Recent epidemiological modeling has provided theoretical evidence of this picture [6]. Within the hosts' immune system trade-offs have also been found, and has in fact recently been measured directly [73, 29].
  (3) Many organs have multiple functions, and the actual structure of such organs are compromises between the various functions they have to fulfil. For instance, mammals have mouths and noses that allow them to efficiently separate food and air while eating, through the use of the secondary palate. Snakes amongst others on the other hand, do not possess a secondary palate and therefore cannot breathe whilst swallowing their prey—a process that may take hours. If one judges the mouth of snakes as mere adaptations for feeding, one might argue that it is much less efficient than the mammalian system. It is of course simply a compromised feeding and breathing system.

In all, it's perhaps a small marvel that natural selection still allows for sufficient adaptability for organisms to adapt to their changing environments.

As a side remark, note that one consequence of the constraints outlined above is the emergence of convergent evolution: if all species in a region are subject to similar constraints then the outcome may often be similar too. Modern car designs are an artificial example (the constraints here are both people's preferences and technological limitations, but perhaps equally important is the search for models that minimize air resistance). A well-known biological example is the form and structure of tropical forest trees. In tropical environments there are few limitations on food and water supplies or temperature. However, because all plants can grow abundantly, competition for light is fierce. In addition, predators and parasites are abundant. So many plant species adopt a similar strategy: grow very fast towards the canopy—yielding slender trunks without branches except at the top—and make very hard wood and thick leaves with pointed ends—thus averting being eaten and avoiding rot from leaves that remain constantly wet.

Chapter 6 focuses on constraints of the third type: trade-offs laws. The examples above are of the general form 'one cannot have too many traits without paying a cost.' The current study looks at the particular case in which there is an active interplay between organisms with multiple possible traits and a parasite that is dependent on this organism. The question then becomes: "With how many traits may one optimally defend oneself against this predator?"

We are concerned both with the specific case of cuckoos and hosts and with the general and fundamental issue of the perfection of evolutionary adaptation. Is natural selection limited in its ability to favour beneficial adaptations? Consider two adaptations that would each enhance the fitness of a single organism (such as the rejection of cuckoo eggs or cuckoo chicks by a host): might such adaptations compete with one another so that the lesser one is lost notwithstanding its value even in the presence of the first?

The parasitic cuckoo

It has been known for centuries that cuckoos do not rear their own young. Instead, they lay their eggs in the nest of a pair of song birds, and after the egg is hatched the foster parents raise it. The female cuckoo removes one of the host's eggs and lays a single egg in the nest. Having laid the egg, the female cuckoo abandons the nest and leaves

the care to the hosts. The egg usually hatches before the eggs of the host. A few hours after hatching, the chick, naked and blind, balances each host egg on its back and ejects it from the nest. Host chicks undergo the same fate and the cuckoo chick becomes the only occupier of the nest. Hence, the parasitized hosts lose all their reproductive success associated with that clutch. The host parents do not intervene in the egg-ejection behaviour of the cuckoo chick. The hosts then rear the single cuckoo chick.

The cuckoo population can be subdivided into so-called *gentes*. On the whole, each gens has specialized in mimicking the colour, patterns of spots and size of the eggs of the particular host it parasitizes [34]. Mimicry is not found, however, in cuckoo chicks when compared with the host's chicks. For instance, in the Reed Warbler *Acrocephalus scirpaceus*, a well studied host of cuckoos, the chicks are small individuals with a pale yellow gape, whereas cuckoo chicks are much larger and have a red gape. This big difference in appearance in cuckoo and host chicks is not exploited by the host to save it from parasitism. The hosts could benefit in two ways if they would reject a cuckoo chick. First, they would save a lot of energy used to raise the cuckoo chick which might, at least, be used by them to survive the winter months. Second, they might be able to immediately start a new clutch.

Absence of nestling discrimination: adaptive or maladaptive?

Understanding the lack of chick rejection behaviour has been hampered by a lack of experimental data: one cannot assess the costs and benefits of a trait that does not exist. Nevertheless, several hypotheses have been put forward. The Evolutionary Lag Hypothesis [119] states that nestling discrimination would be adaptive but not enough time has elapsed for it to evolve. The Evolutionary Equilibrium Hypothesis [88, 89] on the other hand views this trait as maladaptive (a balance between the costs of acceptance and the costs of rejection and recognition errors), and sees the current situation as a stasis.

Parallel to this, there are opposing assumptions about the possible mechanisms of recognition of alien entities in the nest. The discrimination cue may either be learned in early development, or it may be innate. If chick rejection is learned by imprinting, the problem has been solved by Lotem [87], by showing that if hosts make any error in their first attempt to discriminate between their own young and the parasite's, the costs are always greater than the benefits. In the terminology of the two hypotheses the cuckoo-host system is then in evolutionary

equilibrium. Recently, Lawes and Marthews [82] have extended this model to the case of non-evicting parasites and have shown theoretically that parasite chick discrimination may be favoured but only under special circumstances (host nestling survival alongside the parasite is rare, rates of parasitism are high and the average clutch size is large); these are met infrequently.

However, we have not yet been able to explain why chick rejection is nearly absent to explain from a hereditary standpoint. It is this particular question we will address in Chapter 6. In this introduction we will limit ourselves to a general overview of the methods we use to attack this problem, and give an indication what the results are.

Modeling results and implications: defence strategy antagonism and rare enemy effects

Chapter 6 is devoted to the development of a mathematical model to investigate this issue. Its full derivation takes a fair amount of work. Therefore, in this section we will only give the main ingredients of the model.

The model describes the interplay between several types of hosts, (all-accepters, egg-rejecters, chick-rejecters and all-rejecters), and the parasitic cuckoo. Note that an egg-rejecter, for example, does not reject eggs unconditionally, but does so if it is sufficiently sure that the egg it intends to reject is not its own. The model includes both ecological and evolutionary aspects of the problem and is based on an earlier paper by Takasu *et al.* [133]). It incorporates costs incurred by recognition errors but neglects any physiological costs associated with the behavioural capability of rejection. Fitness functions that are dependent on parasitism rate are assigned to each defence strategy. The parasite population is assumed to be limited by host availability, and therefore the parasitism rate in turn depends on the host population and on the defence strategy or strategies that hosts adopt.

The first prediction of the model is that, in almost all circumstances, an equilibrium will be attained where there is only one type of defensive host. As this equilibrium is the result of the competition between different defensive host types, we term this a **Defence Strategy Antagonism**. Here it acts at the population level. It predicts, for example, the absence of a mixed population with some egg-rejecters and some chick-rejecters under most conditions.

The second prediction is that at this equilibrium the fitness of the defensive hosts is the same as the fitness of the non-defensive hosts, i.e.,

those that unconditionally accept parasitic eggs and chicks. In other words, the defensive hosts themselves drive the parasite population down so low that it is only marginally worth continuing such a defence.

Which of the three defensive strategies will the hosts adopt? The model's third prediction is that it is most likely, by far, to be egg rejection. Chick rejection is inherently more costly than egg rejection, because of the delay incurred in putting it into effect and the damage that a cuckoo chick will cause before it is discovered. It will therefore only be favoured over egg rejection if it is much less prone to error, e.g., if mimicry of host eggs by the parasite has led to high failure rates in the egg-rejection process. All-rejection is in most circumstances even less likely to prevail, essentially because of the Rare Enemy Effect [38]. Due to the egg-rejection mechanisms in place, the number of parasitic chicks encountered by a host is reduced. It is important to realize that the mechanisms do not have to be efficient, and the reduction in encounters does not have to be drastic; a small effect is sufficient to tip the balance. In short, this first line of defence may suppress the emergence of the second, which may be thought of as Defence Strategy Antagonism at the individual level. An important exception to this conclusion is that all-rejection might be expected if the chick-rejection strategy is nearly cost-free, i.e., scarcely ever results in the erroneous rejection of a host chick.

So in all, both of the two chick-rejection strategies are favoured if they have very low costs. At the end of the chapter we review two recent reports of chick rejection behaviour [81, 60], and see whether the observed strategies fit with the predictions of the model. The two papers suggest that the defence strategy is indeed innate—making the mathematical model applicable— and that the model properly predicts the conditions under which host birds should reject cuckoo chicks.

CHAPTER 2

Self-contact for rods on cylinders

2.1. Introduction

The study of self-contact in elastic rods has seen some remarkable progress over the last ten years, with highlights such as the numerical work of Tobias, Coleman, and Swigon [136, 28, 27], the introduction of global curvature by Gonzalez and co-workers [58], and the derivation of the Euler-Lagrange equations for energy minimization by Schuricht and Von der Mosel [124]. Parallel advances have been made on the highly related ideal knots and Gehring links, where ropelength is minimized instead of elastic energy [25, 123, 24].

Despite this progress important questions remain open. We are still far from understanding analytically the solutions of the Euler-Lagrange equations for general contact situations. Even if we limit ourselves to global minimizers of an appropriate energy functional, we can prove little about the form of solutions as soon as contact is taken into account.

For instance, a long-standing conjecture for closed elastic rods is that in the limit of long rods under constant twist the global energy minimizer should be a ply (double helix) with a loop on each end. If a structure of this type is assumed, then the limiting pitch angle can be determined [135]; but the difficult part actually consists in showing that global minimizers have this structure. Incidentally, since local minimizers of different type have been found numerically [28, 27], the restriction to global minimizers appears to be essential.

This example is typical for the current state of understanding: if assumptions are made on the set of contacts, then characterizations are possible [47, 131, 135, 66], but for unrestrained geometry little is known

This chapter is based on [68].

rigorously. It shows how our lack of understanding of energy minimizers is intimately linked to the lack of knowledge about structure of the contact set. Examples show that this structure can be non-trivial: for instance, non-contiguous contact appears at the end of a ply in an elastic rod [27].

In this chapter we study a problem of self-contact of an elastic rod in which the rod has reduced freedom of movement: the centreline of the rod is constrained to lie on the surface of a cylinder (Figure 2.1). In contrast to the full three-dimensional problem referred to above, the reduced dimensionality of this problem enables us to give a near-complete characterization of global minimizers, without making any *a priori* assumptions on the structure of the contact set. Notwithstanding this, determining the structure of the contact set is a central element of this study.

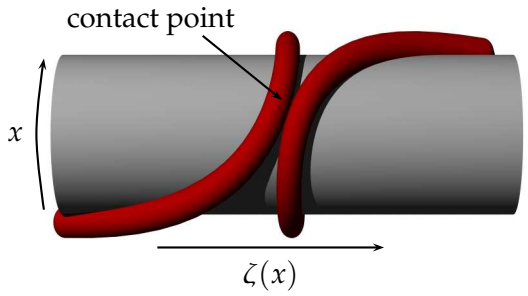


Figure 2.1: The centreline of a rod on a cylinder is described using cylindrical coordinates: the independent variable x is the tangential coordinate, and the position of the centreline is given by the function $\zeta(x)$ measuring distance along the cylinder axis.

We transform the classical Cosserat model of an elastic, unshearable rod of circular cross-section into a more convenient form. The functional that is to be minimized (representing stored energy and work done by the end moment) is

$$F(u) = \int_0^T [\, a(u)u'^2 + b(u) \,],$$

where

$$a(u) = \frac{1}{4\pi^2} \frac{1}{(1+u^2)^{5/2}}, \tag{2.1}$$

$$b(u) = \frac{1}{r^2(1+u^2)^{3/2}} - \frac{2M}{Br} \frac{\sqrt{1+u^2}-u}{\sqrt{1+u^2}}. \tag{2.2}$$

Here r is the radius of the cylinder, M the moment applied to the end of the rod, and B is the bending coefficient of the rod. The centreline of the rod is characterized by $\zeta(x)$, which measures distance along the cylinder axis as a function of a tangential independent variable x. The unknown in this minimization problem is the derivative $u(x) = \zeta'(x)$, which may by thought of as the cotangent of the angle between the centreline tangent and the cylinder axis; u is zero when the rod curls around the cylinder orthogonal to the axis, and $u = \pm\infty$ when the rod is parallel to the axis. This transformation is detailed in Section 2.3.

The most interesting part of the variational problem is the transformed contact condition (condition of non-self-penetration). In this study we take the thickness of the rod to be zero; then the non-self-penetration condition is

$$\int_x^{x+1} u \geq 0 \qquad \text{for all} \quad 0 \leq x \leq T-1, \tag{2.3}$$

where the interval $[x, x+1]$ corresponds to one full turn around the cylinder; this condition formalizes the intuitive idea that non-self-penetration is equivalent to the condition 'that the rod remain on the same side of itself'. This condition on u makes the variational problem a non-local obstacle problem. Non-zero thickness requires a contact condition that is substantially more involved than (2.3); we comment on this situation in Section 2.3.

Both the background in rod theory and the independent mathematical context of this minimization problem raise questions about the solutions:

(1) Do solutions exist?
(2) What is the minimal, and what is the maximal regularity of minimizers?
(3) When is there contact, i.e., when is the contact set

$$\omega_c := \left\{ x \in [0, T-1] \ : \ \int_x^{x+1} u = 0 \right\} \tag{2.4}$$

non-empty?
(4) Given that $\omega_c \neq \emptyset$, what is the structure of ω_c? Is the contact simply contained in a single interval, or is the structure more intricate, as in the examples of contact–skip–contact at the end of a ply [28] and in a (ropelength minimizing) clasp [24]?
(5) What form do the contact forces take?

(6) Does the solution inherit the symmetry of the formulation? This is the case for a symmetric rod on a cylinder without contact condition [65], but need not be true when taking contact into account.

In the rest of this chapter we address these questions.

2.2. Results

The first main result of this investigation (Theorem 2.2) shows that the contact condition (2.3) is essential: without this condition the centreline of a rod will intersect itself. A little experiment with some string wrapped around a pencil will convince the reader that this is the case. We also prove the regularity result that a constrained minimizer u is of class $W^{2,\infty}$, and we derive the Euler-Lagrange equation

$$N(u)(x) := -2a(u(x))u''(x) - a'(u(x))u'^2(x) + b'(u(x)) = \int_{x-1}^{x} f, \tag{2.5}$$

where the Lagrange multiplier f is a non-negative Radon measure with support contained in the contact set ω_c (Theorem 2.5).

From stationarity alone, which is the basis of Theorem 2.5, the characterization of f as a positive Radon measure appears to be optimal; no further information can be extracted. In Section 2.7 we use two different additional assumptions to further characterize the contact set and subsequently the measure f. In both cases we obtain the important result that the contact set is a (possibly empty) interval and that the measure f is a sum of Dirac delta functions, as represented schematically in Figure 2.2. The weighting of the delta functions is shown in the middle of Figure 2.2: there is a linear decrease or increase in weight from one side of the contact set to the other (Theorem 2.17). Since f may be interpreted as the contact force, we deduce that

- The contact force is concentrated in at most two tangential positions x_1 and x_2, and in integer translates of $x_{1,2}$;
- The magnitude of the contact force is maximal at the contact point where the rod lifts off, and decreases linearly with each turn. Figure 2.3 graphically illustrates this behaviour.

The decrease in contact force with each turn can be understood in the following way. The difference between the contact forces on either side of the rod creates a resulting force exerted on the rod, and the two resultant forces that act at $x_{1,2} \bmod 1$ point in opposite directions. If we imagine a single, closed ring with two forces acting on it in this way,

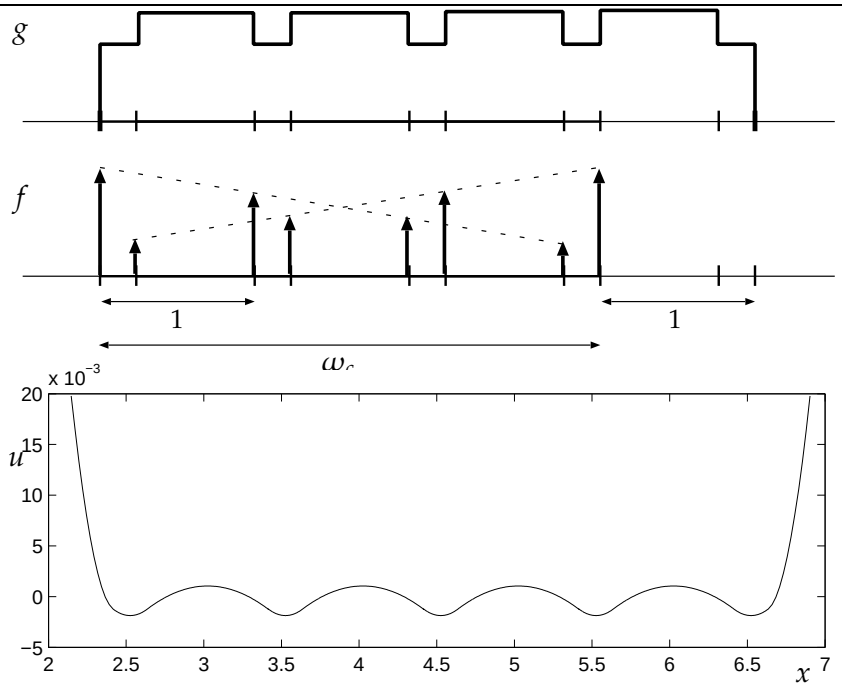


Figure 2.2: The function $g(x) = \int_{x-1}^{x} f$ is piecewise constant (top); the jumps correspond to Dirac delta functions in f (middle). Note that the support of g is the set $\omega_c + [0,1]$ by the definition of g. The solution u corresponding to f and g is shown at the bottom.

the two forces create a moment that will bend the ring. This also happens with the coil of the current problem, as is demonstrated by the small but definite oscillations in the numerical solutions calculated in Section 2.9.

As mentioned above, the crucial result that the contact set is connected requires additional assumptions. If we step back from this rod-on-cylinder model, and allow a and b to be general given functions, then for a large class of such functions the nonlinear operator on the left-hand side of (2.5) $N(u)$ satisfies a version of the comparison principle,

$$Nu_1 \geq Nu_2 \quad \Longrightarrow \quad u_1 \geq u_2,$$

(see Definition 2.19 for the precise statement). For such functions a and b, any stationary point has a connected contact set (Theorem 2.21). The argument is based on the observation that non-contact in some interval (α, β) implies that $f = 0$ on (α, β) and therefore that the right-hand

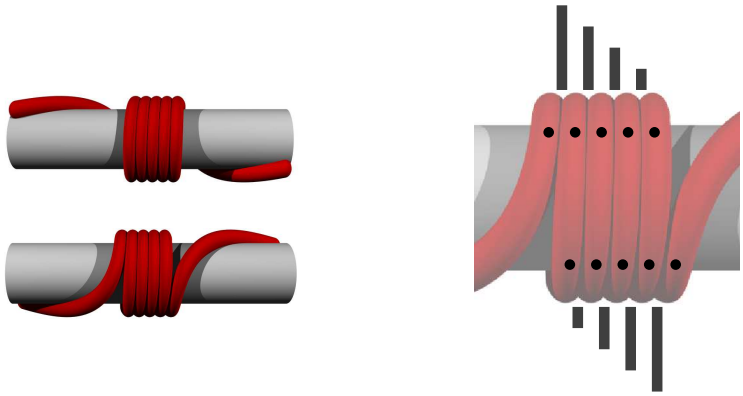


Figure 2.3: A typical rod configuration (left; front and back views) that minimizes energy and satisfies the contact condition. On the right the bars indicate the contact forces corresponding to the arrows in Figure 2.2. (The analysis of this study assumes zero rod thickness—in this picture the rod has been fattened for presentation purposes.)

side of (2.5),

$$g(x) = \int_{x-1}^{x} f, \tag{2.6}$$

is non-increasing on (α, β) and non-decreasing on $(\alpha - 1, \beta - 1)$.

Importantly, however, the functions a and b given in (2.1) are such that the associated operator mostly fails to satisfy this comparison principle. We therefore also take a different approach, in which we obtain the same result by only considering global minimizers. Here we use a different argument, based on constructing other minimizers by cutting and pasting; the combined condition of minimization and non-contact in an interval (α, β) implies the existence of additional regions of non-contact outside of the interval (α, β), implying that the right-hand side of (2.5) is constant on (α, β). From this the result follows (Theorem 2.22).

In both cases, the fact that the contact set is an interval implies that the boundary of the contact set is 'free'—the measure f is zero on an additional interval of length one extending on both ends of ω_c. This implies that the right-hand side g is increasing and decreasing at the same time—except at points that lie at integer distance from the two boundary points. This imposes the specific structure on g and f that is shown in Figure 2.2.

The issue of symmetry of minimizers is a subtle one, which again depends on the presence or absence of a comparison principle. The comparison principle simplifies the structure of solutions: all stationary points are symmetric (up to an unimportant condition on b). Without a comparison principle, and more precisely when minimization of F favours oscillation, this is no longer true, and even stationary points that are global minimizers may be asymmetric (Section 2.8).

Using the characterization of f and g derived earlier we use two numerical methods to investigate constrained minimizers (Section 2.9); one is a method of direct solution, using a boundary-value solver, and the other a continuation method. A typical solution is shown in Figure 2.3.

The simple structure of the functional and the contact condition suggest that the methods and results of this study might be applicable to other systems than this particular rod-on-cylinder model. We therefore state and prove our results for general functions a and b. The main requirements are that a and b are smooth and that a is positive; other conditions are mentioned in the text below.

2.3. Problem setting: derivation of the rod-on-cylinder model

Kinematics

Consider an elastic rod of circular cross-section that is constrained to lie on a cylinder, and which is subject to a force T and a moment M at the ends. We assume that at the rod ends, T, and M are maintained parallel to both the axis of the cylinder and the axis of the rod, but that the loading device leaves the rod ends free to rotate around the circumference of the cylinder; the ends of the rod therefore need not be coaxial. The rod is naturally straight and inextensible, and material cross-sections are assumed to remain orthogonal to the centreline. We will derive a minimization problem for rods of length 2ℓ and later take the limit $\ell \to \infty$.

In the Cosserat rod theory [7, Ch. VIII] the configuration of this rod is characterized by a right-handed orthogonal rod-centred coordinate frame of directors, $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3\}$, each a function of the arclength parameter s. The director $\boldsymbol{d}_3$ is assumed parallel to the centreline tangent, and by the assumption of inextensibility the centreline curve $\boldsymbol{r}$ satisfies

$$\dot{\boldsymbol{r}} = \boldsymbol{d}_3,$$

where the dot denotes differentiation with respect to arclength. The strain of the rod is characterized by the vector-valued function $\boldsymbol{u}$ given by

$$\dot{\boldsymbol{d}}_k = \boldsymbol{u} \times \boldsymbol{d}_k, \qquad k = 1, 2, 3.$$

When decomposed as $\boldsymbol{u} = \kappa_1 \boldsymbol{d}_1 + \kappa_2 \boldsymbol{d}_2 + \tau \boldsymbol{d}_3$, the components may be recognized as the two curvatures $\kappa_{1,2}$ and the twist τ.

We choose a fixed frame of reference $\{\epsilon 1, \epsilon 2, \epsilon 3\}$, where $\epsilon 3$ is parallel to the cylinder axis, and we relate the frame $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3\}$ to this frame by a particular choice of Euler angles $\{\theta, \psi, \phi\}$ [69, 65]. In this parametrization θ is the angle between $\boldsymbol{d}_3$ and $\epsilon 3$ (or between the centreline and the cylinder axis), ψ characterizes the rotation around the cylinder axis, and ϕ is a partial measure of the rotation between cross-sections. The condition that the centreline of the rod lie on the surface of a cylinder of radius r translates into the kinematic condition

$$\dot{\psi} = \frac{1}{r} \sin\theta, \tag{2.7}$$

where the dot denotes differentiation with respect to the arclength coordinate s. Note that it is natural *not* to restrict ψ to an interval of length 2π. In terms of the remaining degrees of freedom $\{\theta, \phi\}$ the curvatures and twist are given by

$$\begin{aligned}
\kappa_1 &= \dot{\theta} \sin\phi - \frac{1}{r} \sin^2\theta \cos\phi, \\
\kappa_2 &= \dot{\theta} \cos\phi + \frac{1}{r} \sin^2\theta \sin\phi, \\
\tau &= \dot{\phi} + \frac{1}{r} \sin\theta \cos\theta.
\end{aligned}$$

Energy, work, and a variational problem

For a given rod the strain energy is given by [69],

$$\begin{aligned}
E(\theta, \tau) &= \frac{B}{2} \int_{-\ell}^{\ell} (\kappa_1^2 + \kappa_2^2) + \frac{C}{2} \int_{-\ell}^{\ell} \tau^2 \\
&= \frac{B}{2} \int_{-\ell}^{\ell} \dot{\theta}^2 + \frac{B}{2r^2} \int_{-\ell}^{\ell} \sin^4\theta + \frac{C}{2} \int_{-\ell}^{\ell} \tau^2.
\end{aligned}$$

Here B and C are the bending and torsional stiffnesses respectively. To determine the work done by the tension and moment at the ends of the rod we need to characterize the generalized displacements associated

with these generalized forces. For the tension T the associated displacement is the shortening S,

$$S(\theta) = \int_{-\ell}^{\ell} (1 - \cos\theta).$$

The generalized displacement associated with the moment M is the end rotation, which is well-defined by the assumption of constant end tangents. It is common to identify the end rotation with a link-like functional

$$L = \int_{-\ell}^{\ell} (\dot{\phi} + \dot{\psi}) = [\phi + \psi]_{-\ell}^{\ell}.$$

As demonstrated in [67], this identification is correct in an open set around the undeformed configuration $\theta \equiv 0$, but loses validity when $|\theta|$ takes values larger than π. Although nothing we have seen suggests that in an energy-minimizing situation θ would take values outside of the admissible interval $(-\pi, \pi)$, we have no rigorous argument to guarantee that θ remains inside that interval, and therefore we are forced to assume this. In terms of the variables θ and τ this functional then takes the form

$$L(\theta, \tau) = \int_{-\ell}^{\ell} \left(\tau + \frac{1}{r} \sin\theta(1 - \cos\theta) \right).$$

Here we assume rigid loading in shortening and dead loading in twist, i.e., we prescribe the shortening S and the moment M, which implies that the tension T and the end rotation L are unknown and to be determined as part of the solution. This loading condition leads to the minimization problem

$$\min \{E(\theta, \tau) - ML(\theta, \tau) : S(\theta) = \sigma\}$$

for given $\sigma > 0$. The tension T has a natural interpretation as a Lagrange multiplier associated with the constraint of S.

We can simplify this minimization problem by first minimizing with respect to τ for fixed θ, from which we find $\tau \equiv M/C$; re-insertion yields the final minimization problem

$$\min \{F(\theta) : S(\theta) = \sigma\} \tag{2.8}$$

with

$$F(\theta) = E(\theta) - ML(\theta) = \frac{B}{2} \int \dot{\theta}^2 + \frac{B}{2r^2} \int \sin^4\theta - \frac{M}{r} \int \sin\theta(1 - \cos\theta). \tag{2.9}$$

We are interested in localized forms of deformation, in which the deformation is concentrated on a small part of the rod and in which boundary effects are to be avoided, and therefore we take an infinitely long rod and consider θ, F, and S to be defined on the whole real line and assume $\theta \to 0$ as $|s| \to \infty$.

Behaviour of minimizers

The Euler-Lagrange equations associated with the minimization problem (2.8) can be written as a Hamiltonian system with one degree of freedom,

$$\frac{1}{2}\dot{\theta}^2 + V(\theta) = H, \tag{2.10}$$

for a particular V. In this system two independent parameters remain, which may be interpreted as a scaled cylinder radius $\tilde{r} = rM/B$ and a combined loading parameter $m = M/\sqrt{BT}$.

Solutions of the original minimization problem are orbits of this Hamiltonian system that are homoclinic to zero, and such orbits have been studied in detail in [65]. Among the findings are

(1) For all values of $\tilde{r}$ ranges of m exist with orbits that are homoclinic to the origin;
(2) At some parameter points these homoclinic orbits 'collide' with saddle points. The saddle points correspond to helical solutions (constant angle θ) and close to these collisions the homoclinic orbit has a large region of near-constant angle θ.

In Figure 2.4 a bifurcation diagram is shown with two such collisions, one at a forward helix ($0 < \theta < \pi/2$, at $m = m_{c_2}$) and one at a backward helix ($\pi/2 < \theta < \pi$, at $m = m_{c_1}$).

In [65] the question of stability of these solutions, both local and global, was left untouched. If we interpret the combined load parameter m as a (reciprocal) tension T (with the moment M fixed) then the nature of the bifurcation diagram in Figure 2.4, involving as it does the mechanically conjugate variables S and T, suggests that in each peak the right curve is locally stable [92]. With two peaks occurring however, this does not allow us to predict where the globally stable solution is located.

In this study we focus on global energy minimization. Corollary 2.3 below states that for sufficiently large shortening, and when contact effects are neglected, *global energy minimizers always intersect themselves*. It is this result that forms the main motivation of our analysis: since

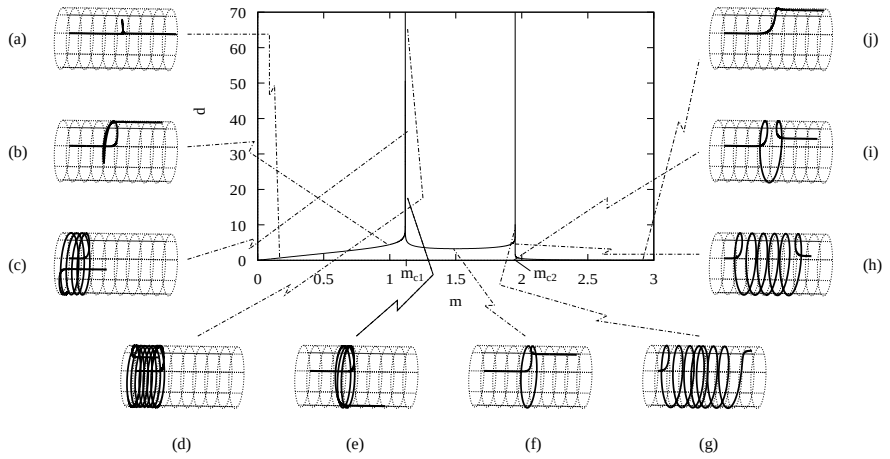


Figure 2.4: A load-displacement diagram showing shortening $d = SM/B$ of stationary points as a function of the (combined) load m (from [65]). Contact effects are not taken into account. The peaks divide this diagram into three sections. The solutions in the middle section intersect themselves, whilst the solutions on the right do not. The section on the left consists of heteroclinic connections between $\theta = 0$ and $\theta = 2\pi$ which are not considered here. For sufficiently large shortening, the rod configuration that has lowest energy is on the self-penetrating branch, as shown by Corollary 2.3.

energy minimization without appropriate penalization leads to self-intersection, the non-self-intersection condition is necessary for physically acceptable solutions.

Translation to (u, ψ)-coordinates

To study the case in which self-contact is taken into account, it is necessary to properly restrict the class of admissible functions in the minimisation problem (2.8). In three dimensions a variety of different descriptions of self-contact exists for rods of finite thickness, each with subtle advantages and disadvantages (see e.g., the introduction of [58]). For a rod on a cylinder the situation is simpler, since the freedom of movement is essentially two-dimensional—similar to that of a curve in a plane. We focus on rods of zero thickness, and implement non-self-penetration as non-self-intersection of the centreline. In terms of the

unknown $\theta(\cdot)$ as introduced above, this condition can be written as

$$z(s_1) - z(s_2) \neq 0 \quad \text{for all} \quad s_1 \neq s_2 \text{ with } \quad \psi(s_1) - \psi(s_2) = 0 \mod 2\pi, \tag{2.11}$$

where we have used the previous equation (2.7) for ψ and the axial coordinate z:

$$\dot{\psi} = \frac{1}{r}\sin\theta, \qquad \dot{z} = \cos\theta.$$

We now make the assumption that z can be written as a function of ψ, or, equivalently, that ψ is monotonic along the rod. This assumption is satisfied for solutions of the problem without contact having $\theta < \pi$, as given by equation (2.10). If we include a contact condition of the form (2.11), then we are unable to prove that ψ is monotonic, and in fact it is conceivable that this monotonicity is only valid for *global* energy minimizers.

Under the assumption that z can be written as a function of ψ, we introduce a dimensionless axial coordinate $\zeta = z/r$, and write $'$ for differentiation with respect to ψ. The functional F in (2.9) then transforms to

$$F(\zeta) = \frac{Br}{2}\int_0^T \frac{\zeta''^2}{(1+\zeta'^2)^{\frac{5}{2}}} + \frac{B}{2r}\int_0^T \frac{1}{(1+\zeta'^2)^{\frac{3}{2}}} - M\int_0^T \frac{\sqrt{1+\zeta'^2}-\zeta'}{\sqrt{1+\zeta'^2}},$$

with shortening

$$S(\zeta) = r\int_0^T \left[\sqrt{1+\zeta'^2} - \zeta'\right].$$

Here $[0, T]$, the domain of definition of ψ, is *a priori* unknown, since the ends of the rod are free to move around the cylinder.

In these variables non-self-intersection is easily characterized. Since ψ is monotonic, let us assume it to be increasing (this amounts to an assumption on the sign of the applied moment M). Admissible functions are defined by the following condition:

$$\forall \psi \in [0, T - 2\pi] : \zeta(\psi + 2\pi) - \zeta(\psi) \geq 0. \tag{2.12}$$

Note that it is only necessary to rule out self-intersection after a *single* turn; if contact exists after multiple turns, contact also exists (potentially elsewhere) after a single turn.

The contact condition (2.12) is the novel part in this variational problem. Here we focus on the effect that this condition has on the minimization problem, and therefore simplify by

- fixing the domain size T, and accordingly removing the shortening constraint;

- replacing the mechanically correct boundary conditions $\zeta' = \infty$ by a more convenient condition $\zeta' = 1$.

In terms of the new variables $x = \psi/2\pi$ and $u(x) = \zeta'(\psi)$ we recover the problem of the introduction.

These boundary conditions can be described as follows. By prescribing $\zeta' = u = 1$ at the ends of the rod we fix the angle between the rod and the centreline to $\pi/4$. By removing the shortening constraint we allow the ends of the rod to move freely in the axial direction; in contrast, the fixing of T prevents the rod ends from moving tangentially. We believe that these changes have little effect on that part of the rod that is implicated in the contact problem; but this is a topic of current research.

Zero thickness

The assumption of zero rod thickness can not be relaxed without introducing important changes in the formulation (see Figure 2.5). At thickness ϵ, the distance in the ζ-direction between two parallel consecutive centrelines in contact is $\epsilon/\sin\theta$, where θ is the angle between the centrelines and the cylinder axis. Therefore non-zero thickness can not be introduced by simply replacing the right-hand side in (2.12) by ϵ; the angle of the centrelines is to be taken into account, implying that the right-hand side of (2.12) will depend on ζ'.

To make matters worse, when the centrelines are not parallel, i.e., when $u = \zeta'$ is not constant, the minimal-distance connection between two consecutive turns depends on values of ζ' nearby (see [101] for a thorough treatment of the geometry of this issue); it is not clear whether for the present case of a rod on a cylinder any simpler impenetrability condition can be found than the well-known global curvature condition [58].

2.4. Existence and the contact condition

In this section we state precisely the problem under discussion and show that minimizers exist. We also study the minimization problem *without* the contact constraint, and show that minimizers will intersect themselves.

Let $U = 1 + X$, where $X = H_0^1(0, T)$, and $Y = C([0, T-1])$. Let the functional $F : U \to \mathbb{R}$ be defined as in the Introduction,

$$F(u) = \int_0^T [\, a(u)u'^2 + b(u)\,],$$

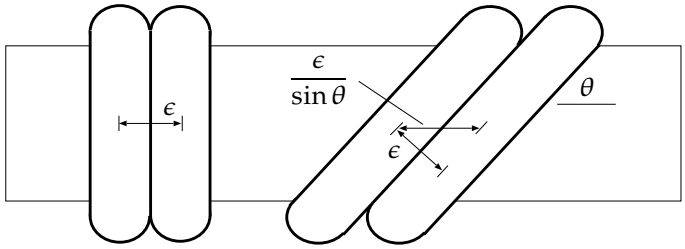


Figure 2.5: Two configurations of a rod of thickness ϵ. This illustrates that for rods with positive thickness one cannot simply replace the contact condition $Bu \geq 0$ by $Bu \geq \epsilon$; a more involved condition is necessary.

and introduce the constraint operator $B : U \to Y$ given by

$$B(u)(x) = \int_x^{x+1} u.$$

With the set of admissible functions given as

$$K := \{u \in U : \ B(u)(x) \geq 0 \ \forall x \in [0, T-1]\}$$

the central problem is

Problem (A): Find a function $u^* \in U$ such that

$$F(u^*) = \min\{F(u) : \ u \in K\}.$$

We first prove existence of minimizers for Problem (A).

LEMMA 2.1. *Let $T > 0$. Assume that $a(u) \geq a_0 > 0$, and that $b(u)$ is Lipschitz continuous. Then there exists $u^* \in K$ such that*

$$F(u^*) = \min\{ F(u) \ : \ u \in K \}.$$

PROOF. Let $\{u_n\} \subset K \subset U$ be a minimizing sequence. We first prove that $\int b(u_n)$ is bounded from below.

Since minimization of F is equivalent to minimization of $F - Tb(0)$, we can assume without loss of generality that $b(0) = 0$. Using the Lipschitz continuity of b and the Poincaré inequality we have

$$\|b(u_n)\|_{L^1} \leq c\|u_n\|_{L^1} \leq c_1(T + \|u_n - 1\|_{L^1}) \leq c(1 + \|u_n'\|_{L^2}).$$

Here and below c is a possibly changing constant that does not depend on n. Then

$$\begin{aligned}\int b(u_n) &\geq -c(1+\|u_n'\|_{L^2}) \\ &\geq -c\left(1+\frac{1}{a_0}\Big(\int a(u_n){u_n'}^2\Big)^{\frac{1}{2}}\right) \\ &\geq -c\left(1+\frac{1}{a_0}\Big(F(u_n)-\int b(u_n)\Big)^{\frac{1}{2}}\right) \\ &\geq -c\left(1+\frac{1}{a_0}\Big(c-\int b(u_n)\Big)^{\frac{1}{2}}\right).\end{aligned}$$

Hence

$$\int b(u_n) \geq -D \tag{2.13}$$

for a suitable constant D.

Together with the boundedness of $F(u_n)$, (2.13) implies that u_n is bounded in X. Hence $\{u_n\}$ contains a subsequence $\{u_{n_m}\}$ that converges weakly in X to a limit u^*. Since F is lower semicontinuous with respect to weak convergence,

$$F(u^*) \leq \liminf_{m\to\infty} F(u_{n_m}),$$

implying that u^* is a minimizer. □

As we mentioned in Section 2.3, if contact is not taken into account—if F is minimized in U rather than in the smaller set K—condition. In the theorem below we actually prove a stronger statement. We write F_T instead of F to indicate explicitly the dependence on the interval $[0,T]$.

THEOREM 2.2 (Minimization without contact condition). *Assume that a and b are of class C^1, and that a is strictly positive. Assume that some $\bar{u} < 1$ exists such that*

$$-\infty < \inf_{\mathbb{R}} b < \inf_{u\geq\bar{u}} b. \tag{2.14}$$

For each $T > 0$, let u_T be a minimizer corresponding to the minimization problem on domain $[0,T]$,

$$\min\{F_T(u) \ : \ u \in U\}. \tag{2.15}$$

Then there exists $c > 0$ independent of T such that

$$|\{x \in [0,T] : u_T(x) \geq \bar{u}\}| \leq c\left(1+\sqrt{T}\right).$$

The function b given in (2.1) achieves its minimum at $u = -\infty$, regardless of the value of Mr/B; therefore it satisfies the condition (2.14) for every $\bar{u} < 1$.

COROLLARY 2.3 (Minimizers violate the contact condition). *In addition to the conditions of Theorem 2.2, assume that $\bar{u} < 0$. If T is sufficiently large, then $B(u_T)(x) < 0$ for some $x \in [0, T-1]$.*

PROOF OF THEOREM 2.2. We first use a standard argument to give an upper bound on the energy $F_T(u_T)$. Choose a T-dependent constant $\underline{u}_T < \bar{u}$ such that

$$b(\underline{u}_T) < \inf_{u \geq \bar{u}} b(u) \qquad \text{and} \qquad 0 < b(\underline{u}_T) - \inf_{\mathbb{R}} b \leq T^{-1/2}.$$

For any $T \in \mathbb{R}^+$ we construct a new continuous symmetric function $\tilde{u}_T \in U$ such that $\tilde{u}_T = \underline{u}_T$ on $[1, T-1]$, and $F_T(\tilde{u}_T) \leq C + Tb(\underline{u})$, where C does not depend on T. Since u_T minimizes F_T, it also follows that

$$F_T(u_T) \leq F_T(\tilde{u}_T) \leq C + Tb(\underline{u}_T). \tag{2.16}$$

Among other things this inequality implies that for large T a minimizer u_T can not be the constant function 1.

The Euler-Lagrange equation associated with this minimization problem is

$$-2a(u)u'' - a'(u)u'^2 + b'(u) = 0, \tag{2.17}$$

which can also be written as a one-degree-of-freedom Hamiltonian system

$$-a(u)u'^2 + b(u) = H. \tag{2.18}$$

It follows that for any minimizer u,

(1) $b(u(x)) = H$ at any stationary point x of u;
(2) $b(u(x)) \geq H$ for all $x \in [0, T]$;
(3) $b(1) > H$.

The third statement follows from noting that if $b(1) = H$ then $u \equiv 1$ would be the unique solution of (2.18).

We now show that any minimizer u is bi-monotonic, i.e., increasing or decreasing away from a minimum or maximum. Suppose instead that u has two internal stationary points, a minimum at x_1 and a maximum at x_2; assume for definiteness that $0 < x_1 < x_2 < T$. Note that $u(x_1) < 1 < u(x_2)$, since the solution of the Hamiltonian system is a periodic orbit oscillating between the values $u(x_1)$ and $u(x_2)$; the inequality $u(x_1) < 1 < u(x_2)$ follows from the boundary condition. Now pick a point $x_{12} \in (x_1, x_2)$ such that $u(x_{12}) = 1$.

Construct a new function

$$\tilde{u}(x) = \begin{cases} u(x) & 0 \le x \le x_1 \\ u(x_1) & x_1 \le x \le x_1 + T - x_{12} \\ u(x - T + x_{12}) & x_1 + T - x_{12} \le x \le T \end{cases}$$

Then

$$\begin{aligned} F_T(\tilde{u}) &= \int_0^{x_1} [\, a(u)u'^2 + b(u)\,] + \int_{x_1}^{x_1+T-x_{12}} b(u(x_1)) \\ &\quad + \int_{x_1}^{x_{12}} [\, a(u)u'^2 + b(u)\,] \\ &= \int_0^{x_{12}} [\, a(u)u'^2 + b(u)\,] + \int_{x_1}^{x_1+T-x_{12}} b(u(x_1)) \\ &= \int_0^{x_{12}} [\, a(u)u'^2 + b(u)\,] + H(T - x_{12}) \\ &< F_T(u). \end{aligned}$$

Therefore the assumption of two stationary points is contradicted. Note that by (2.18) the solution also is symmetric in $[0, T]$.

We now return to the sequence of functions u_T. Setting $A = \{x \in [0, T] : u_T(x) \ge \bar{u}\}$ we have

$$\begin{aligned} C + Tb(\underline{u}_T) &\ge F_T(u_T) \\ &\ge |A| \inf_{u \ge \bar{u}} b(u) + (T - |A|) \inf_{\mathbb{R}} b \\ &= |A| \left(\inf_{u \ge \bar{u}} b(u) - \inf_{\mathbb{R}} b \right) + T \inf_{\mathbb{R}} b, \end{aligned}$$

so that

$$|A| \left(\inf_{u \ge \bar{u}} b(u) - \inf_{\mathbb{R}} b \right) \le C + T\big(b(\underline{u}_T) - \inf_{\mathbb{R}} b\big) \le c\,(1 + \sqrt{T}).$$

This concludes the proof. □

REMARK 2.4. By a very similar argument one may show the following statement: if $\min_{\mathbb{R}} b$ is uniquely achieved at some $\bar{u} \in \mathbb{R}$, then

$$\|u_T - \bar{u}\|_{L^\infty\left(\sqrt{T}, T - \sqrt{T}\right)} \longrightarrow 0 \qquad \text{as} \qquad T \longrightarrow \infty.$$

2.5. The Euler-Lagrange equation

We characterize the duality (X, X') by identifying the smooth functions on $[0, T]$ with a dense subset of X' via the duality pairing

$$_{X'}\langle \xi, x\rangle_X = \int_0^T \xi x.$$

Similarly we identify Y' with the space of Radon measures $RM([0, T-1])$ via the same duality pairing, defined for smooth functions,

$$_{Y'}\langle \eta, y\rangle_Y = \int_0^T \eta y.$$

Where necessary, we extend Radon measures in Y' by zero outside of their domain $[0, T-1]$.

THEOREM 2.5. *Assume that a and b are globally Lipschitz continuous, and that $a \geq a_0 > 0$. Let $u \in U$ be a solution of Problem (A). Then $u \in W^{2,\infty}(0,T)$ and there exists a Radon measure $f \in Y'$ such that*

$$-2a(u(x))u''(x) - a'(u(x))u'^2(x) + b'(u(x)) = \int_{x-1}^{x} f(s)\,ds \qquad (2.19)$$

for almost every $x \in (0,T)$. Moreover $f \geq 0$ and $\operatorname{supp} f \subset \omega_c$.

DEFINITION 2.6. *A function $u \in U$ is called a stationary point if it there exists a Radon measure $f \in Y'$, with $f \geq 0$ and $\operatorname{supp} f \subset \omega_c$, such that (2.19) is satisfied.*

In the rest of the chapter we will often drop the arguments in (2.19) and write

$$-2a(u)u'' - a'(u)u'^2 + b'(u) = \int_{x-1}^{x} f.$$

The proof of Theorem 2.5 follows along the lines of [11]. We fix the function u, with contact set ω_c defined in (2.4), and introduce the cone of admissible perturbations V,

$$V := \{v \in X \,:\, \exists \{\varepsilon_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+, \varepsilon_n \to 0 \text{ such that } B(u+\varepsilon_n v) \geq 0 \,\forall n \in \mathbb{N}\}.$$

LEMMA 2.7. *Let u be a minimizer. Then $F'(u)\cdot v \geq 0$ for all $v \in \overline{V}$.*

PROOF. For any $v \in V$, the fact u is a minimizer implies that

$$F(u+\varepsilon_n v) - F(u) \geq 0 \qquad \text{for all } n \in \mathbb{N}.$$

The conditions on a and b imply that F is Fréchet differentiable in u (this follows from the conditions on a and inspection of (2.21) below), so that

$$0 \leq F(u+\varepsilon_n v) - F(u) = \varepsilon_n F'(u)\cdot v + o(\varepsilon_n ||v||_X),$$

from which it follows that $F'(u) \cdot v \geq 0$. Now, given any $v \in \overline{V}$, take a sequence $v_m \subset V$ that converges to v in X. Since $F'(u) : X \to \mathbb{R}$ is a continuous linear operator, $F'(u) \cdot v_m \to F'(u) \cdot v$. Hence $F'(u) \cdot v \geq 0$ for any $v \in \overline{V}$. □

$\overline{V}$ can be characterized in a more convenient way:

LEMMA 2.8. *For any $u \in K$,*

$$\overline{V} = W := \{v \in X \ : \ Bv \geq 0 \text{ on } \omega_c\}.$$

We postpone the proof to the end of this section.

$\overline{V}$ is a closed convex cone, with dual cone

$$\overline{V}^\perp = \{\gamma \in X' : \ \langle \gamma, v \rangle \geq 0 \quad \forall v \in \overline{V}\}.$$

Let

$$P = \{y \in Y : \ y \geq 0 \text{ on } \omega_c\}.$$

This also is a closed convex cone, with dual cone

$$P^\perp = \{f \in Y' : \ \langle f, y \rangle \geq 0 \quad \forall y \in P\}.$$

LEMMA 2.9. *If $f \in P^\perp$, then* supp $f \subset \omega_c$ *and $f \geq 0$.*

PROOF. Given any y with support in ω_c^c, $y \in P$ and $-y \in P$. Hence $\langle f, y \rangle = 0$ and therefore supp $f \subset \omega_c$. Now take $y \in Y$ positive. Then in particular $y \geq 0$ on ω_c, and $y \in P$. By definition of $P^\perp$ this implies $f \geq 0$. □

We now use the following lemma to characterize $\overline{V}^\perp$ in a different way.

LEMMA 2.10. *Let Y be a Banach space, and $P \subset Y$ a closed convex cone with dual cone $P^\perp$. Let X be a second Banach space, and $A : X \to Y$ a bounded linear operator. Let K be the following cone in X:*

$$K = \{u \in X : \ Au \in P\}.$$

Then the dual cone $K^\perp$ can be characterized by

$$K^\perp = \{A^T g \in X' : g \in P^\perp\}.$$

The proof of this lemma can be found in [11]. An immediate consequence of Lemma 2.10 is

COROLLARY 2.11.

$$\overline{V}^\perp = \{B^T f \in X' : f \in P^\perp\}.$$

We now turn to the proof of the main theorem of this section.

PROOF OF THEOREM 2.5. We have seen that, since u is a minimizer, $F'(u) \in \overline{V}^{\perp}$ and

$$\overline{V} = \{v \in X \,:\, B(v) \geq 0 \text{ on } \omega_c\},$$

by Lemmas 2.7 and 2.8. By Corollary 2.11 there exists an $f \in P^{\perp}$ such that $F'(u) = B^T f$, and by Lemma 2.9 $\operatorname{supp} f \subset \omega_c$ and $f \geq 0$. The conjugate operator B^T is easily seen to be given by

$$B^T \phi(x) = \int_{x-1}^{x} \phi(s)\, ds \tag{2.20}$$

for a smooth function $\phi \in Y'$, where ϕ is implicitly extended by zero outside of the interval $[0, T-1]$. We use the same notation for a general Radon measure $f \in Y'$.

Lastly, direct computation gives

$$F'(u) \cdot v = \int_0^T \left[\, 2a(u)u'v' + a'(u)u'^2 v + b'(u)v \,\right], \tag{2.21}$$

and hence we obtain the equation

$$-2[a(u(x))u'(x)]' + a'(u(x))u'(x)^2 + b'(u(x)) = \int_{x-1}^{x} f \tag{2.22}$$

in the sense of distributions.

We now turn to the statement of regularity. Since $f \in RM([0, T-1])$, the function g (defined in (2.6)) is uniformly bounded. Since all terms in (2.22) except the first are in $L^1(0,T)$, we have $a(u)u' \in W^{1,1}$, and the lower bound on a implies that $u \in W^{2,1}(0,T)$. Since $W^{2,1} \subset W^{1,\infty}$, the second term is now known to be in L^∞, and again the lower bound on a is used to obtain $u \in W^{2,\infty}(0,T)$. This regularity of u implies that the distributional equation (2.22) is also satisfied almost everywhere. □

We still owe the reader the proof of Lemma 2.8.

PROOF OF LEMMA 2.8. $\overline{V} \subset W$: Since $B : X \to Y$ is continuous, W is closed, and therefore it suffices to show that $V \subset W$. Take any $v \in V$ and $x \in \omega_c$. Then $B(u + \varepsilon_n v)(x) \geq 0$, and by definition of ω_c, $Bu(x) = 0$, implying that $Bv(x) \geq 0$. It follows that $v \in W$.

$W \subset \overline{V}$: First consider $w \in W$ such that $\operatorname{supp}(Bw)_-$ (the support of the negative part of Bw) is contained in ω_c^c. We claim that $w \in V$, for

which we have to show that there exists

$$\{\varepsilon_n\} \subset \mathbb{R}^+,\ \varepsilon_n \to 0, \text{ such that } B(u+\varepsilon_n w) \geq 0\ \forall n \in \mathbb{N}.$$

For $x \in \omega_c$, $Bu(x) = 0$, and since $Bw(x) \geq 0$ we have $B(u+\varepsilon_n w)(x) \geq 0$ for any sequence $\{\varepsilon_n\} \subset \mathbb{R}^+$. For the complement ω_c^c, note that since $\operatorname{supp}(Bw)_-$ is compact and contained in the open set ω_c^c, there exists $\delta > 0$ such that $Bu \geq \delta > 0$ on $\operatorname{supp}(Bw)_-$. Hence, if $\varepsilon_n \leq \delta \|Bw\|_{L^\infty}^{-1}$, then $Bu + \varepsilon_n Bw \geq 0$ on $\operatorname{supp}(Bw)_-$. Note that $Bu \geq 0$ on $[0, T-1]$, and $Bw \geq 0$ on $(\operatorname{supp}(Bw)_-)^c$. This means that $Bu + \varepsilon_n Bw \geq 0$ on ω_c^c. Together with $Bw \geq 0$ on ω_c, this implies $w \in V$.

Finally, consider a general $w \in W$. Fix a smooth function $\phi \in X$ with $\phi > 0$ on $(0, T)$; note that $B\phi \geq c > 0$. We approximate w by the function $w_\varepsilon := w + \varepsilon\phi$. We claim that $\operatorname{supp}(Bw_\varepsilon)_- \subset \omega_c^c$ for sufficiently small $\varepsilon > 0$. It then follows that $w_\varepsilon \in V$ and $w_\varepsilon \to w$, implying that $w \in \overline{V}$.

To prove the claim, note that $w \in X \subset L^\infty$. Hence Bw is Lipschitz continuous, with Lipschitz constant $2||w||_{L^\infty}$. Hence, for small enough ε,

$$\begin{aligned} Bw_\varepsilon(x) &= Bw(x) + \varepsilon B\phi(x) \\ &\geq Bw(y) + \varepsilon B\phi(y) - 3||w||_{L^\infty}|x-y|. \end{aligned} \tag{2.23}$$

Suppose $Bw_\varepsilon(x) < 0$ and $y \in \omega_c$. Then $Bw(y) \geq 0$, and by (2.23),

$$-\varepsilon B\phi(y) > -3||w||_{L^\infty}|x-y|,$$

or

$$|x-y| > \frac{\varepsilon B\phi(y)}{3||w||_{L^\infty}}.$$

Therefore $d(\operatorname{supp}(Bw)_-, \omega_c) > C\varepsilon$ for a suitable $C > 0$. Hence $\operatorname{supp}(Bw_\varepsilon)_- \subset \omega_c^c$ for small enough ε, which proves the claim. □

2.6. Characterization of stationary points

For this section we assume that the conditions of Theorem 2.5 are met.

LEMMA 2.12. *Let u be a stationary point, and let g be defined as in (2.6).*

(1) *For all $x \in \omega_c$, $u(x) = u(x+1)$ and $u'(x) \leq u'(x+1)$.*
(2) *If ω_c contains an interval $[x_0, x_1]$, then*
   - *$u'(x) = u'(x+1)$ for all $x \in (x_0, x_1)$;*
   - *$u''(x) = u''(x+1)$ and $g(x) = g(x+1)$ for almost all $x \in (x_0, x_1)$.*

This lemma imposes an interesting form of periodicity on the solution and the right-hand side g. Although the constraint is a non-local one, on an interval of contact of length L the solution actually only has the degrees of freedom of an interval of length one; the other values follow from this assertion.

PROOF. Since $x \in \omega_c$,

$$\int_x^{x+1} u = 0.$$

Hence, since $Bu(x) = \int_x^{x+1} u \in W^{3,\infty}$, and $Bu \geq 0$,

$$0 = \frac{d}{dx} \int_x^{x+1} u = u(x+1) - u(x),$$

and

$$0 \leq \frac{d^2}{dx^2} \int_x^{x+1} u = u'(x+1) - u'(x).$$

If $Bu = 0$ on $[x_0, x_1]$, then the inequality above becomes an equality a.e. on the interior (x_0, x_1), implying that

$$u'(x) = u'(x+1) \text{ on } (x_0, x_1).$$

The periodicity of u'' and g now follow from (2.19). □

LEMMA 2.13. *Let u be a stationary point, and assume that ω_c contains an interval I. Then*

$$\int_x^{x+1} g$$

is constant on Int I.

PROOF. By Lemma 2.12 $u(x) = u(x+1)$ for all $x \in I$, and $u''(x) = u''(x+1)$ a.e. on I. In addition, $u'(x) = u'(x+1)$ for all $x \in \operatorname{Int} I$. Hence

$$x \mapsto \int_x^{x+1} \left[-2a(u)u'' - a'(u)u'^2 + b'(u)\right] \tag{2.24}$$

is constant on I. But by (2.19), (2.24) is equal to

$$\int_x^{x+1} \int_{s-1}^{s} f = \int_x^{x+1} g.$$

□

The following two lemmas and the theorem that follows are essential in determining the structure of the right-hand side g and therefore of the measure f. The main argument is the following. The function g has no reason to be monotonic; its derivative in x equals

$f(x) - f(x-1)$, and although f is a positive measure this difference may be of either sign. However, if for instance a left end point x_0 of ω_c is flanked by a non-contact interval $(x_0 - 1, x_0)$, then the measure f is zero on that interval, and the function g is non-decreasing on $(x_0, x_0 + 1)$. It is this argument, repeated from both sides, that allows us to determine completely the structure of the function g and the underlying measure f.

Notation Let $[x_0, x_1] \subset \omega_c$. Define

$$p \equiv x_1 - x_0 \pmod 1, \tag{2.25}$$

and

$$P = \min\{n \in \mathbb{N} : n \geq x_1 - x_0\}. \tag{2.26}$$

Throughout the rest of this chapter τ is the translation operator defined by

$$(\tau u)(x) = u(x+1). \tag{2.27}$$

LEMMA 2.14. *Let u be a stationary point, such that ω_c contains an interval $[x_0, x_1]$. Assume furthermore that*

$$\operatorname{supp} f \cap (x_0 - 1, x_0) = \emptyset. \tag{2.28}$$

Then

(1) *if $x_1 - x_0 \in \mathbb{N}$, then g does not decrease on each of the subintervals*

$$(x_0 + i, x_0 + i + 1),\ i = 0, 1, \ldots, P;$$

(2) *if $x_1 - x_0 \notin \mathbb{N}$, then g does not decrease on each of the subintervals*

$$(x_0 + i, x_0 + i + 1),\ i = 0, 1, \ldots, P - 1,$$

nor does it on

$$(x_0 + P, x_1 + 1).$$

PROOF. On $(x_0, x_0 + 1)$,

$$g' = f - \tau^{-1} f \overset{(2.28)}{=} f \geq 0,$$

and therefore g is non-decreasing on $(x_0, x_0 + 1)$. By Lemma 2.12, $g(x) = g(x+1)$ for almost all $x \in (x_0, x_1)$. This implies that on each consecutive interval $(x_0 + i, x_0 + i + 1)$, $i = 1, \ldots, P - 1$, g does not decrease. By the same reasoning, if $x_1 - x_0 \in \mathbb{N}$, then this also holds for $(x_0 + P, x_0 + P + 1) = (x_1, x_1 + 1)$. If not, then it holds for $(x_0 + P, x_1 + 1)$. □

REMARK 2.15. Let u be a stationary point. Define the mirror image $v(x) = u(T - x)$, and $h(x) = f(T - x - 1)$. Then (v, h) solves

$$\begin{cases} -2a(v)v'' - a'(v)v'^2 + b'(v) = \int_{x-1}^{x} h, \\ v(0) = v(T) = 1, \end{cases}$$

and hence is also a stationary point.

Applying Lemma 2.14 to (v, h) yields for (u, f):

LEMMA 2.16. *Let u be a stationary point such that ω_c contains an interval $[x_0, x_1]$. Assume furthermore that*

$$\operatorname{supp} f \cap (x_1 + 1, x_1 + 2) = \emptyset.$$

(1) *if $x_1 - x_0 \in \mathbb{N}$, then g does not increase on each of the subintervals*

$$(x_0 + i, x_0 + i + 1),\ i = 0, 1, \dots, P;$$

(2) *if $x_1 - x_0 \notin \mathbb{N}$, then g does not increase on each of the subintervals*

$$(x_0 + p + i, x_0 + p + i + 1),\ i = 0, 1, \dots, P - 1,$$

nor does it on

$$(x_0, x_0 + p).$$

To combine the previous two lemmas, let

$$\begin{aligned} X_i &= x_0 + i, \qquad i = 0, \dots, P, \\ Y_i &= x_0 + p + i, \quad i = 0, \dots, P. \end{aligned} \tag{2.29}$$

THEOREM 2.17. *Let u be a stationary point such that the contact set ω_c contains an interval $[x_0, x_1]$. Suppose that*

$$\operatorname{supp} f \cap \big\{(x_0 - 1, x_0) \cup (x_1 + 1, x_1 + 2)\big\} = \emptyset.$$

Then there exists $G \in \mathbb{R}$ such that

(1) *if $x_1 - x_0 \in \mathbb{N}$, then $g \equiv G$ on $(x_0, x_1 + 1)$, and*

$$f|_{(x_0-1, x_1+2)} = G \sum_{i=0}^{P} \delta(x - X_i).$$

(2) *if $x_1 - x_0 \notin \mathbb{N}$, then*

$$g(x) = \begin{cases} g_1 & \text{on } [X_i, Y_i], \quad i = 0, \dots, P, \\ g_2 & \text{on } [Y_i, X_{i+1}], \quad i = 0, \dots, P - 1, \end{cases} \tag{2.30}$$

and

$$f|_{(x_0-1, x_1+2)} = \sum_{i=0}^{P-1} a_i \delta(x - X_i) + b_i \delta(x - Y_i),$$

where $a_i = (G - \frac{i}{P})g_1$ *and* $b_i = \frac{G+i}{P}g_1$, *and*

$$g_1 := \frac{GP}{P+1-p} \in \Big(\frac{PG}{P+1}, G\Big), \tag{2.31}$$

$$g_2 := \frac{G(P+1)}{P+1-p} = \frac{P+1}{P}g_1 \in \Big(G, \frac{(P+1)G}{P}\Big). \tag{2.32}$$

PROOF. (1) $x_1 - x_0 \in \mathbb{N}$.

By Lemma 2.14, g does not decrease on the intervals (X_i, X_{i+1}), $i = 0, 1, \ldots, P$, and by Lemma 2.16 g does not increase on these intervals either. Hence g is constant on each interval. By Lemma 2.13 the constant is the same on each interval, i.e., that $g \equiv G$ on $(x_0, x_1 + 1)$. This also implies that within the interval $(x_0 - 1, x_1 + 2)$, f can only have support in the points $x_0 = X_0, X_1, \ldots, X_P = x_1$, yielding the formula for f in the statement of the theorem.

(2) $x_1 - x_0 \notin \mathbb{N}$.

Combining Lemma 2.14 and 2.16, we find that g is constant on each interval (X_i, Y_i) and (Y_i, X_{i+1}), $i = 0, 1, \ldots, P-1$, and on (X_P, Y_P). By Lemma 2.12, $g(x) = g(x+1)$ for almost all $x \in (x_0, x_1)$, and hence g takes three values, 0, and g_1 and g_2 (say) on $(x_0, x_1 + 1)$. We choose $g = g_1$ on (X_i, Y_i), $i = 0, 1, \ldots, P$, and $g = g_2$ on the intervals inbetween, (Y_i, X_{i+1}), $i = 0, 1, \ldots, P-1$; outside of the interval (x_0, x_1), g vanishes. By Lemma 2.13,

$$G = \int_x^{x+1} g = \int_x^{x+p} g_1 + \int_{x+p}^{x+1} g_2 = pg_1 + (1-p)g_2. \tag{2.33}$$

Either $g_1 = g_2 = G$ or $g_1 < G < g_2$. The first case implies that $g \equiv G$ on $[x_0, x_1 + 1]$. This implies that f does not only have support in X_0, X_1, …, X_{P-1}, but by reasoning for the mirror image (v, h) it also implies that f has support in $Y_0, Y_1, \ldots, Y_{P-1} = x_1$. This is impossible. Hence $g_1 < G < g_2$. The support of f on $(x_0 - 1, x_1 + 2)$ is now seen be to limited to the set given in the statement of the theorem.

Thus we conclude that f is a sum of delta functions, but we still have to determine the weights a_i and b_i. Since $f = 0$ on $(x_0 - 1, x_0)$, we have $g_1 = g(x_0+) = f(x_0)$. Here we abuse notation, and write $f(x)$ for the weight of the Dirac delta function at x. Now we have the following recurrence relations:

$$\begin{aligned} f(X_i) + f(Y_i) &= g_2, \\ f(Y_i) + f(X_{i+1}) &= g_1, \end{aligned}$$

for $i = 0, 1, \dots, P-1$. Solving this system we obtain

$$\begin{aligned} f(X_i) &= f(X_0) - i(g_2 - g_1) = g_1 - i(g_2 - g_1), \\ f(Y_i) &= (i+1)(g_2 - g_1). \end{aligned}$$

In addition, since $x_1 = Y_{P-1}$, $f(x_1) = P(g_2 - g_1)$. On the other hand, $g_1 = h(T - x_1 - 1) = f(x_1)$. This implies

$$g_2 = \frac{P+1}{P} g_1.$$

To conclude,

$$a_i = f(X_i) = \big(G - \frac{i}{P}\big) g_1,$$

and

$$b_i = f(Y_i) = \frac{i+1}{P} g_1.$$

By Lemma 2.13,

$$p g_1 + (1-p)\frac{P+1}{P} g_1 = G,$$

which yields

$$p = P\Big(1 - \frac{G}{g_1}\Big) + 1.$$

Solving for g_1 now yields all required results. □

As we will see in the next section, the contact set of u is connected in many important cases. Hence Theorem 2.17 allows us to give concise expressions for g in cases that ω_c is an interval of positive length (using the Heaviside function H):

COROLLARY 2.18. *If the contact set is an interval of positive length, then g equals the explicit function*

$$g(x; x_0, x_1, G) = \begin{cases} H(x - x_0) - H(x - x_1 - 1) & \text{if } x_1 - x_0 \in \mathbb{N}, \\ g_1\big(H(x - x_0) - H(x - x_1 - 1)\big) + & \\ \quad + (g_2 - g_1)\sum_{i=1}^{P}\big[H(x - X_i) - H(x - Y_i)\big] & \text{if } x_1 - x_0 \notin \mathbb{N}. \end{cases} \tag{2.34}$$

Here the coefficients $g_{1,2}$ are computed from x_0, x_1, and G by (2.25), (2.26), (2.31). and (2.32).

Figure 2.6 shows examples of both cases. For the remaining two cases of a stationary point that has a single or no contact point, g is immediately clear: with a single contact point,

$$g(x) = \begin{cases} m & \text{on } [x_0, x_0 + 1], \\ 0 & \text{otherwise,} \end{cases}$$

for a suitable constant $m \geq 0$, while when there is no contact then obviously $g \equiv 0$.

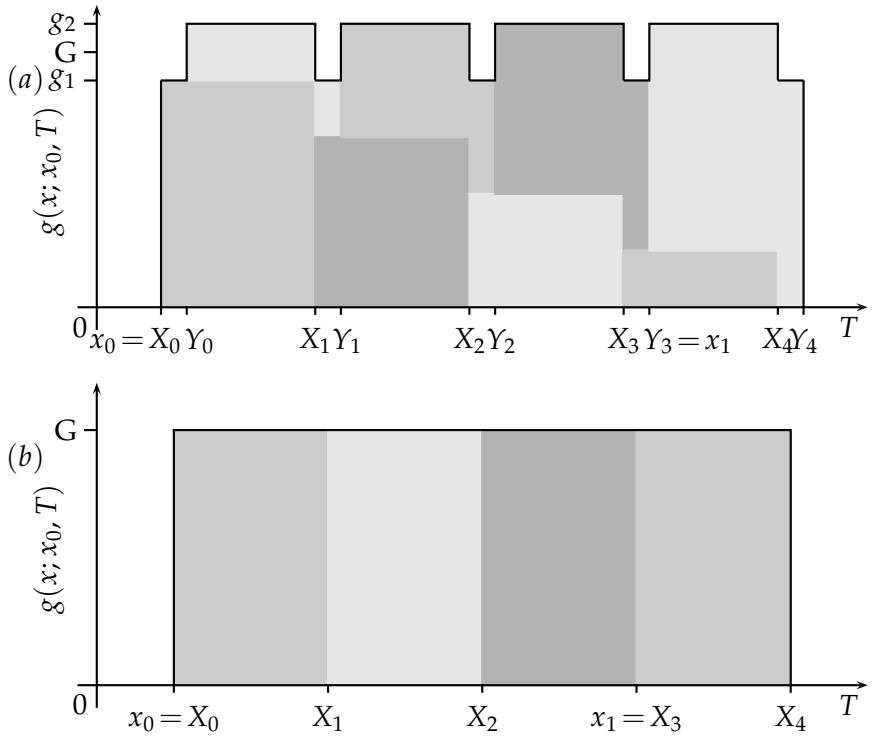


Figure 2.6: A generic picture of $g(x; x_0, T)$ for $T - x_0 \notin \mathbb{N}$ (a) and for $T - x_0 \in \mathbb{N}$ (b). The light gray shaded areas represent the contributions of the individual delta functions of the corresponding Radon measures f. As an example, in (b) f consists of four Dirac deltas, all with mass G, at $x_0 = X_0, x_0 + 1 = X_1, \ldots, x_0 + 3 = X_3$.

2.7. The contact set is an interval

In order to extract more information on the right-hand side g and the measure f than that given by Theorem 2.5 we study two cases. In the first case we assume that the operator given by the left-hand side in (2.19) satisfies a version of the classical comparison principle. In the second case we restrict ourselves to global minimizers.

DEFINITION 2.19. *Let N be a (non)linear operator on U. N is said to satisfy the comparison principle if for any $[x_0, x_1] \subset [0, T]$,*

$$\left.\begin{array}{l} Nu_1 \leq Nu_2, \\ u_1(x_0) \leq u_2(x_0), \\ u_1(x_1) \leq u_2(x_1), \end{array}\right\} \implies u_1 \leq u_2 \text{ on } [x_0, x_1].$$

See e.g., [55] or [110] for a general exposition. Operators of the type considered here, i.e.,

$$Nu := -2a(u)u'' - a'(u)u'^2 + b'(u),$$

may fail to satisfy the comparison principle for two reasons. First, the zero-order term $b'(u)$ need not be increasing in u; for instance, the operator $u \mapsto -u'' - u$ does not satisfy the comparison principle on any interval of length 2π or more. In a slightly more subtle manner, the prefactor $a(u)$ of the second-order derivative may also invalidate the comparison principle; see e.g., [55, Section 10.3] for an example.

We conjecture that the 'true' rod functions a and b given in (2.1) do not give rise to a comparison principle: b' is not monotonic, suggesting that on sufficiently large intervals the principle will fail.

We first prove a lemma that will be used in both cases.

LEMMA 2.20. *Let u be a stationary point such that $x_1, x_2 \in \omega_c$. Assume that*

$$(x_1, x_2) \cap \omega_c = \emptyset.$$

Then

$$\int_{x_1}^{x_2} u = \int_{x_1+1}^{x_2+1} u, \tag{2.35}$$

and for any $m \in (0, x_2 - x_1)$,

$$\int_{x_1}^{x_1+m} u < \int_{x_1+1}^{x_1+1+m} u \qquad \text{and} \qquad \int_{x_2-m}^{x_2} u > \int_{x_2-m+1}^{x_2+1} u. \tag{2.36}$$

PROOF. Since $(x_1, x_2) \cap \omega_c = \emptyset$,

$$\int_{x_1+m}^{x_1+m+1} u > 0,$$

for all $m \in (0, x_2 - x_1)$, which implies

$$\begin{aligned}\int_{x_1+1}^{x_1+m+1} u - \int_{x_1}^{x_1+m} u &= \int_{x_1}^{x_1+1} u + \int_{x_1+1}^{x_1+m+1} u - \int_{x_1}^{x_1+m} u \\ &= \int_{x_1+m}^{x_1+m+1} u > 0.\end{aligned}$$

The other two assertions are handled similarly. □

THEOREM 2.21. *Let u be a stationary point, and assume that*

$$Nu := -2a(u)u'' - a'(u)u'^2 + b'(u) \tag{2.37}$$

satisfies the comparison principle. Then ω_c is connected.

PROOF. We proceed by contradiction. Since ω_c is closed, non-connectedness implies the existence of $x_1, x_2 \in \omega_c$ such that $(x_1, x_2) \cap \omega_c = \emptyset$.

Set $v = u - \tau u$. Then $v(x_1) = v(x_2) = 0$ by Lemma 2.12, $\int_{x_1}^{x_2} v = 0$ by (2.35), and

$$\int_{x_1}^{x_1+m} v < 0 \text{ for all } 0 < m < x_2 - x_1 \tag{2.38}$$

by (2.36). Hence there exists an $\bar{x} \in (x_1, x_2)$ such that $v(\bar{x}) = 0$.

From $(x_1, x_2) \cap \omega_c = \emptyset$ it follows that $\operatorname{supp} f \cap (x_1, x_2) = \emptyset$. Hence $g = \int_{x-1}^{x} f$ is a decreasing function on (x_1, x_2) and τg is an increasing function on this interval by previous arguments. Hence $g - \tau g$ is a decreasing function on (x_1, x_2). There are three possibilities, each leading to a contradiction with the comparison principle.

Case 1: $g \geq \tau g$ on (x_1, x_2). On $(x_1, \bar{x})$,

$$\begin{cases} Nu = g \geq \tau g = N\tau u, \\ u(x_1) = \tau u(x_1), \\ u(\bar{x}) = \tau u(\bar{x}). \end{cases}$$

By the comparison principle, $u \geq \tau u$ on $(x_1, \bar{x})$, i.e., $v \geq 0$. But this contradicts (2.38).

Case 2: there exists an $\tilde{x}$ such that $g \geq \tau g$ on $(x_1, \tilde{x})$ and $g \leq \tau g$ on $(\tilde{x}, x_2)$. If $\tilde{x} \geq x_1$, the same argument applies. If $\tilde{x} < \bar{x}$, we consider $(\bar{x}, x_2)$ instead, and apply the same argument. Now we conclude $v \leq 0$ on $(\bar{x}, x_2)$. But observe that from $\int_{x_1}^{\bar{x}} v < 0$ by (2.38) and $\int_{x_1}^{x_2} v = 0$ we have $\int_{\bar{x}}^{x_2} v > 0$, which again implies a contradiction.

Case 3: $g \leq \tau g$ on (x_1, x_2). Again we obtain a contradiction from considering the interval $(\bar{x}, x_2)$. □

For the second case we limit ourselves to global minimizers. The results of this theorem do apply to the functions a and b given in (2.1).

THEOREM 2.22. *Let u be a minimizer. Assume that a and b are of class C^1 and that a is strictly positive. Then ω_c is connected.*

PROOF. As in the proof of Theorem 2.21 we assume that there exist $x_1, x_2 \in \omega_c$ with $(x_1, x_2) \cap \omega_c = \emptyset$ to force a contradiction. Then

$$\operatorname{supp} f \cap (x_1, x_2) = \emptyset, \tag{2.39}$$

and hence g is a decreasing function on (x_1, x_2), and an increasing function on $(x_1 + 1, x_2 + 1)$. Now consider the following two new functions

$$v(x) = \begin{cases} u(x) & \text{on } [0, x_1], \\ u(x+1) & \text{on } [x_1, x_2], \\ u(x) & \text{on } [x_2, T], \end{cases}$$

and

$$w(x) = \begin{cases} u(x) & \text{on } [0, x_1 + 1], \\ u(x-1) & \text{on } [x_1 + 1, x_2 + 1], \\ u(x) & \text{on } [x_2 + 1, T]. \end{cases}$$

Both are admissible, i.e., $v, w \in K$: they are continuous by Lemma 2.12, implying that $v, w \in X$, and the fact that $Bv, Bw \geq 0$ follows from Lemma 2.20. In fact we need certain strict inequalities, which we derive after introducing some notation.

The functions v and w are minimizers. To show this, write

$$F(u|_{[x_1,x_2]}) = \int_{x_1}^{x_2} \left[a(u)u'^2 + b(u) \right].$$

Then since u is a minimizer, and since u and v only differ on $[x_1, x_2]$,

$$F(u|_{[x_1,x_2]}) \leq F(v|_{[x_1,x_2]}) = F(u|_{[x_1+1,x_2+1]}),$$

and similarly

$$F(u|_{[x_1+1,x_2+1]}) \leq F(w|_{[x_1+1,x_2+1]}) = F(u|_{[x_1,x_2]}).$$

This implies that

$$F(u|_{[x_1,x_2]}) = F(u|_{[x_1+1,x_2+1]}),$$

and that $F(u) = F(v) = F(w)$. Every minimizer is also a stationary point, and hence for v and w there exist positive Radon measures f_v and f_w such that $\operatorname{supp} f_v \subset \omega_c(v)$ and $\operatorname{supp} f_w \subset \omega_c(w)$. We also denote $g_v(x) = \int_{x-1}^{x} f_v$ and $g_w(x) = \int_{x-1}^{x} f_w$.

For any $x \in (x_1, x_2)$,

$$\int_x^{x+1} u > 0. \tag{2.40}$$

Let first $x_2 \geq x_1 + 1$. Then for any $x \in (x_1 - 1, x_1)$,

$$\begin{aligned}\int_x^{x+1} v &= \int_x^{x_1} v + \int_{x_1}^{x+1} v \\ &= \int_x^{x_1} u + \int_{x_1+1}^{x+2} u \\ &> \int_x^{x_1} u + \int_{x_1}^{x+1} u \qquad \text{by Lemma 2.20} \\ &= \int_x^{x+1} u \geq 0.\end{aligned}$$

For any $x \in (x_1, x_2 - 1)$ the same is true:

$$\int_x^{x+1} v = \int_{x+1}^{x+2} u > 0,$$

since $x + 1 < x_2$, which allows us to use (2.40). Now let $x_2 < x_1 + 1$. Then for any $x \in (x_1 - 1, x_2 - 1)$, we can repeat the first argument above to conclude

$$\int_x^{x+1} v > 0.$$

Combining these statements we find

$$\int_x^{x+1} v > 0 \text{ for all } x \in (x_1 - 1, x_2 - 1),$$

which implies $\omega_c(v) \cap (x_1 - 1, x_2 - 1) = \emptyset$. Hence $\operatorname{supp} f_v \cap (x_1 - 1, x_2 - 1) = \emptyset$. But since u and v coincide on $[0, x_1]$, we have $g_v|_{[0,x_1]} = g_u|_{[0,x_1]}$, so that $\operatorname{supp} f_u \cap (x_1 - 1, x_2 - 1) = \emptyset$. Combined with (2.39), this implies that $g_u|_{[x_1,x_2]}$ is constant. By symmetry the same is true for $g_u|_{[x_1+1,x_2+1]}$. Note that if $x_2 > x_1 + 1$, then the overlap implies that the two constants are the same; for the other case we now prove this.

Define $z = u - \tau u$; the function z solves the equation

$$\begin{aligned}-2a(u)z'' = g_u - \tau g_u + \{a'(u)u'^2 - a'(\tau u)(\tau u)'^2\} \\ - \{b'(u) - b'(\tau u)\} + \{2a(u) - a(\tau u)\}(\tau u)''\end{aligned} \tag{2.41}$$

on the interval (x_1, x_2). Of the right-hand side, we have seen above that the term $g_u - \tau g_u$ is constant on (x_1, x_2); let us suppose it non-zero for the purpose of contradiction. The function z is of class C^1, and both z and z' vanish at $x = x_{1,2}$. Therefore the assumed regularity on a and b

implies that the expressions between braces are continuous on $[x_1, x_2]$ and zero at $x = x_{1,2}$. The sign of the right-hand side of (2.41) is therefore determined by $g_u - \tau g_u$, and most importantly, is the same at both ends x_1 and x_2; therefore the sign of z, at $x = x_1+$ and $x = x_2-$, is also the same. This contradicts the following consequence of Lemma 2.20:

$$\int_{x_1}^{x_1+m} z < 0 \qquad \text{and} \qquad \int_{x_2-m}^{x_2} z > 0 \qquad \text{for all} \quad 0 < m < x_2 - x_1.$$

This leaves $g_u = \tau g_u$ on (x_1, x_2). But then, by uniqueness of the initial-value problem, $u(x) = u(x+1)$ for all $x \in [x_1, x_2]$, and $[x_1, x_2] \subset \omega_c$, contrary to our assumption that $(x_1, x_2) \cap \omega_c = \emptyset$. □

2.8. Symmetry

In the introduction we raised the question whether the stationary points or minimizers inherit the symmetry of the formulation, or to put it differently, whether non-symmetric solutions exist.

For the discussion of this question it is useful to introduce an equivalent formulation of the Euler-Lagrange equation (2.19) similar to the Hamiltonian-systems formulation used in the proof of Theorem 2.2. For the length of this section we assume that Theorem 2.17 applies and therefore that there is a single contact interval $[x_0, x_1]$.

By multiplying (2.19) with u' and integrating one finds that the function H, defined by

$$H := -a(u)u'^2 + b(u) - gu, \tag{2.42}$$

is piecewise constant, and that H and g jump at the same values of x. The function g takes three values on $[0, T]$, these being the values g_1 and g_2 introduced in Theorem 2.17, and the value $g_0 = 0$ outside of the extended contact interval $[x_0, x_1 + 1]$. (Note that g_1 and g_2 may be equal). We claim that H also takes three values, H_0, H_1, and H_2, and that these values correspond to those of g, i.e., that the *pair* (g, H) takes three values $(0, H_0)$, (g_1, H_1), and (g_2, H_2) (although it may happen that $(g_1, H_1) = (g_2, H_2)$).

To prove this claim, first consider the case of $p > 0$, where p is defined as in (2.25). Then

$$u|_{(x_0,x_0+p)} \equiv u|_{(x_0+1,x_0+1+p)} \qquad \text{and} \qquad g|_{(x_0,x_0+p)} \equiv g|_{(x_0+1,x_0+1+p)}$$

by Lemma 2.12 and (2.30). Therefore H is the same on these two intervals. Repeating this argument for all subintervals of $[x_0, x_1 + 1]$ of the form $(x_0 + k, x_0 + k + p)$ and $(x_0 + k + p, x_0 + k + 1)$ we find that H

takes two values on the interval $[x_0, x_1 + 1]$, H_1 and H_2, and that these coincide with the values g_1 and g_2 of g.

When $p = 0$, a similar argument yields that H takes only one value on $[x_0, x_1 + 1]$ (as does g).

A consequence of this characterization of H is the following lemma:

LEMMA 2.23. *Under the conditions and notation of Theorem 2.17,*

$$u(x_0) = u(x_0 + p) = u(x_0 + 1) = u(x_0 + 1 + p) = \cdots = u(x_1 + 1).$$

PROOF. When $p = 0$ the statement follows from Lemma 2.12. For $p > 0$, note that at any of the *interior* jump points, i.e., at all jump points except x_0 and $x_1 + 1$, we have $[H] = -[g]u$ where $[H] = \pm(H_2 - H_1)$ and $[g] = \pm(g_2 - g_1)$. Regardless of the sign this equation has only one solution u. For the remaining two points x_0 and $x_1 + 1$ the result follows from Lemma 2.12. □

We still need to show that the value of H is the same on both sides of the extended contact interval $[x_0, x_1 + 1]$, so that we can define the value H_0 unambiguously. If one of the ends of this interval equals 0 or T there is nothing to prove; we therefore assume that $\min\{x_0, T - x_1 - 1\} \geq d > 0$. Now multiply (2.19) with the function

$$v(x) = \begin{cases} \frac{x}{d}u'(x) & 0 < x < d \\ u'(x) & d \leq x \leq T - d \\ \frac{T-x}{d}u'(x) & T - d < x < T, \end{cases}$$

and integrate to find

$$-\frac{1}{d}\int_0^d H + \frac{1}{d}\int_{T-d}^T H = 0.$$

Since H is constant on $(0, d)$ and on $(T - d, T)$ the two constant values are equal; we then define H_0 to be this value.

We now turn to the implications of this characterization of solutions (u, g) and the associated pseudo-Hamiltonian function H.

THEOREM 2.24. *Let u be a stationary point with a single contact interval $[x_0, x_1]$. Let p be given as in (2.25), and define the set of jump points $J = \{x_0, x_0 + p, x_0 + 1, x_0 + 1 + p, \ldots, x_1 + 1\}$.*

(1) *There exists $\alpha \in \mathbb{R}$ such that at any $x \in J$, $u'(x) = \pm\alpha$.*

Now assume that b is non-decreasing on $[1, \infty)$.

(2) *If the operator N given in (2.37) satisfies the comparison principle, then u is symmetric on $[0, T]$.*

(3) *If u is a minimizer with $u'(x_0) = -u'(x_1 + 1)$, then u is symmetric on $[0, T]$.*

PROOF. For the first part write

$$u'^2 = \frac{b(u) - gu - H}{a(u)},$$

and note that by the proof of Lemma 2.23 the sum $gu + H$ is continuous.

For the second part, note that by Lemma 2.23 u has the same value on each end of the interval $[x_0, x_0 + p]$ (if $p > 0$) or $[x_0, x_0 + 1]$ (if $p = 0$). By the uniqueness that follows from the comparison principle the function u is symmetric on this interval. By repeating this argument over all subintervals of $[x_0, x_1 + 1]$ we find that u is symmetric on $[x_0, x_1 + 1]$.

The functions $u_1(t) := u(x_0 - t)$ and $u_2(t) := u(x_1 + 1 + t)$, therefore, have the same zeroth and first derivatives at $t = 0$; they satisfy the same equation (2.42) (note that H is symmetric on $[x_0, x_1 + 1]$); therefore the two functions are equal as long as they both exist. This implies that lack of symmetry must stem from a difference in domain of definition of u_1 and u_2 for $t > 0$.

We claim that neither u_1 nor u_2 has an interior maximum. Assuming this claim, the assertion of the theorem follows since the monotonicity of $u_{1,2}$ then implies that the boundary condition $u_{1,2}(t) = 1$ has at most one solution t.

Now assume that u_1 has a maximum at $t_1 > 0$. The function u_1 is solution of the Hamiltonian system (2.42), where H and g are constant for $t > 0$. As in the proof of Theorem 2.2, therefore $u_1(t_1) > 1$. Choose a bounded interval $I \subset [0, \infty)$ such that $u > 1$ on $\operatorname{Int} I$ and $u(\partial I) = 1$.

The reduced functional $\tilde{F}(v) = \int_I [a(v)v'^2 + b(v)]$ has a global minimizer $\tilde{v}$ in the class of functions v satisfying $v(\partial I) = 1$. From studying the perturbation $v \mapsto \min\{v, 1\}$ and using the monotonicity of b it follows that $\tilde{v} \leq 1$ on I. By the comparison principle this is the only stationary point of $\tilde{F}$, a conclusion that contradicts the fact that u_1 is a different stationary point.

For the third part, first note that the support of the continuous function $x \mapsto \int_x^{x+1} u$ is $[x_0, x_1]$; therefore

> for every $\epsilon > 0$ there exists $\delta > 0$ such that any perturbation v with $d(\operatorname{supp} v, [x_0, x_1 + 1]) > \epsilon$ is admissible provided $\|v\|_{L^\infty} < \delta$.

We will use this below.

The assumption on the derivatives places us in the same position as above: the functions $u_1(t) := u(x_0 - t)$ and $u_2(t) := u(x_1 + 1 + t)$ are equal as long as they both exist. Again we will show that neither may have an interior maximum, but by a different argument.

Assume that u_1 has a maximum. By defining $t_1 = x_0$ the boundary condition on u takes the form $u_1(t_1) = 1$. Pick

$$\max\{1, u_1(0)\} < \beta < \max\{u_1(t) : 0 \leq t \leq t_1\}$$

and define the set $S = \{t \in [0, t_1] : u_1(t) \geq \beta\}$; we can assume that for $\epsilon = \inf S > 0$ we have $\max\{u_1(t) : 0 \leq t \leq t_1\} - \beta < \delta$ for the associated δ given above.

Now define $v(t) = \min\{\beta, u_1(t)\}$. The function v is admissible by construction; it differs from u_1 only on the set S, and therefore the difference in energy is given by (with a slight abuse of notation)

$$F(v) - F(u_1) = \int_S \left[-a(u_1)u_1'^2 + b(\beta) - b(u_1)\right] < 0.$$

This contradicts the assumption of minimality. □

The conditions of Theorem 2.24 are quite sharp. We demonstrate this with two examples.

Example 1: b is decreasing on $[1, \infty)$. It is relatively straightforward to construct a non-symmetric stationary point by choosing an appropriate function b that is decreasing on $[1, \infty)$, thus showing that part 2 of Theorem 2.24 is sharp.

Take a symmetric stationary point u for which $u \leq 1$ on $[0, T]$, $u'(T) > 0$, and for which the contact set is bounded away from $x = T$ (see the next section for examples). Close to $x = T$, the function u satisfies

$$u'^2 = \frac{b(u) - H}{a(u)}$$

for some $H \in \mathbb{R}$, and since $u'(T) > 0$, $b(1) > H$. Now change $b(u)$ for $u > 1$ such as to have (for instance) $b(2) = H$, and continue the solution u past $x = T$. By construction $u(T + \tilde{T}) = 2$, for some $\tilde{T} > 0$, and $u'(T + \tilde{T}) = 0$; by symmetry then $u(T + 2\tilde{T}) = 1$. The new function u defined on the domain $[0, t + 2\tilde{T}]$ is a non-symmetric stationary point (Figure 2.7).

Example 2: equal (non-opposite) derivatives on $\partial\omega_c$. For certain functions b and domains $[0, T]$ global minimization favours breaking of

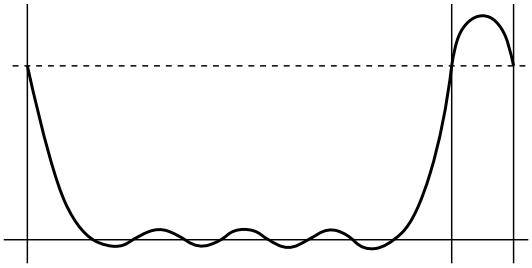


Figure 2.7: A non-symmetric stationary point can be constructed by defining $b(u)$ appropriately for $u > 1$.

symmetry. We demonstrate this for the functional

$$F(u) = \int \left[u'^2 + \alpha(1-u^2)^2 \right],$$

where α will be chosen appropriately. We consider the functional F on functions $u : [0,1] \to \mathbb{R}$ with boundary conditions $u(0) = u(1) = 0$; although this is slightly different from the setup in the rest of the study, it simplifies the argument, and the extension to a more general situation is intuitively clear.

We will show that

$$\inf \left\{ F(u) : \int u \geq 0 \right\} \; < \; \inf \left\{ F(u) : \int u \geq 0 \text{ and } u \text{ is symmetric} \right\}. \tag{2.43}$$

The infimum on the right-hand side is bounded from below,

$$F(u) \geq (1 - \alpha c/2) \int u'^2 + \alpha,$$

by the Poincaré inequality

$$\int_0^d u^2 \leq cd^2 \int_0^d u'^2 \qquad \text{for all } u \text{ with } u(0) = 0 \text{ and } \int u = 0. \tag{2.44}$$

The function $v(x) = a + \cos(b(1 - x/d))$ is optimal in this inequality, where $a \simeq 0.22$ and $b \simeq 4.49$ are determined by the boundary condition $v(0) = 0$ and the integral condition $\int v = 0$. The Poincaré constant equals $c \simeq 0.0495$. Note that for symmetric functions u we may take $d = 1/2$.

At the function $w(x) = \sin 2\pi x$ the functional F has the value $F(w) = 2\pi^2 + 3\alpha/8$. For all $\alpha \in (16\pi^2/5, 2/c] \simeq (31.6, 40.3]$ therefore

$$F(w) = 2\pi^2 + 3\alpha/8 < \alpha \leq \inf\left\{F(u) : \int u \geq 0 \text{ and } u \text{ is symmetric}\right\},$$

which demonstrates (2.43).

The reason for this preference for asymmetry can be recognized in the constant in the Poincaré inequality (2.44) (see Figure 2.8). For symmetric functions the relevant class is $\{u : [0, 1/2] \to \mathbb{R} : u(0) = \int u = 0\}$, and for more general functions $\{u : [0, 1] \to \mathbb{R} : u(0) = u(1) = \int u = 0\}$. For this latter class the Poincaré coefficient is achieved by the function w above with the value $\overline{c} = 1/4\pi^2 \simeq 0.0253$, which is larger than $c(1/2)^2 = 0.0124$.

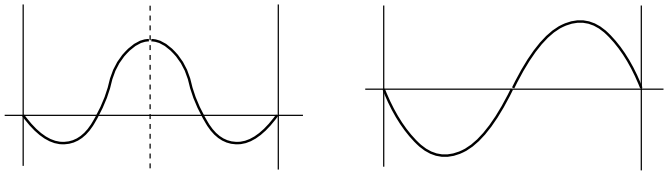


Figure 2.8: Under symmetry conditions the effective domain, the domain on which $\int u = 0$, is half the actual domain size. Equivalently, more (costly) oscillations are necessary.

2.9. Numerical simulations

In this section we describe in detail our numerical simulations of stationary points of F under constraint, i.e., of solutions of

$$-2a(u)u'' - a'(u)u'^2 + b'(u) = \int_{x-1}^{x} f, \tag{2.45}$$

$$u(0) = u(T) = 1, \tag{2.46}$$

$$\operatorname{supp} f \subset \omega_c, \tag{2.47}$$

$$f \text{ a positive Radon measure}, \tag{2.48}$$

$$\int_x^{x+1} u \geq 0 \quad \forall x \in [0, T-1]. \tag{2.49}$$

We concentrate on the case in which the solution is symmetric and the contact set is non-empty, and we use the fact that the right-hand side in the differential equation can be characterised explicitly (see (2.34)). We

further simplify by replacing the inequality (2.49) by the condition that the function $x \mapsto \int_x^{x+1} u$ has a second-degree zero at $x = x_0$, leading to the new system in the unknowns (u, x_0, G)

$$-2a(u)u'' - a'(u)u'^2 + b'(u) = g(x; x_0, T - x_0 - 1, G), \tag{2.50}$$

$$u(0) = u(T) = 1, \tag{2.51}$$

$$u(x_0) = u(x_0 + 1), \tag{2.52}$$

$$\int_{x_0}^{x_0+1} u = 0. \tag{2.53}$$

For brevity we shall write $\gamma(x; x_0, T, G)$ for $g(x; x_0, T - x_0 - 1, G)$.

LEMMA 2.25. *Assume that the operator on the left-hand side of (2.50) satisfies the comparison principle. Then any solution of problem (2.45–2.49) with non-empty contact set is also a solution of (2.50–2.53); vice versa, any solution of (2.50–2.53) is also a solution of (2.45–2.49).*

PROOF. Since the implication (2.45–2.49) $\Longrightarrow$ (2.50–2.53) follows by construction, it suffices to show the opposite implication; in fact, since an admissible measure f can be constructed from any $\gamma(x; x_0, T, G)$, it is sufficient to show that solutions of (2.50–2.53) satisfy

$$\int_x^{x+1} u \geq 0 \qquad \forall x \in [0, T - 1].$$

We show slightly more, namely that

$$\int_x^{x+1} u = 0 \qquad \forall x \in [x_0, T - x_0 - 1]$$

and that

$$\int_x^{x+1} u > 0 \qquad \forall x \notin [x_0, T - x_0 - 1].$$

The function u is symmetric by Theorem 2.24. Since $u(x_0) = u(x_0 + 1)$,

$$u(x_0) = u(x_0 + 1) = u(T - x_0) = u(T - x_0 - 1) =: \bar{u}.$$

Set $u_1(x) = u(x_0 + x)$, and $u_2(x) = u(x_0 + x + 1)$ for all $x \in [0, T - 2x_0 - 1]$. By construction, $\gamma(x; x_0, T, G) = \gamma(x + 1; x_0, T, G)$ for all $x \in [x_0, T - x_0 - 1]$. Hence, if we set $h(x) = \gamma(x + x_0; x_0, T, G)$, for all $x \in [0, T - 2x_0 - 1]$, then u_1 and u_2 both satisfy

$$-2a(v)v'' - a'(v)v'^2 + b'(v) = h,$$
$$v(0) = v(T - 2x_0 - 1) = \bar{u},$$

By uniqueness, $u_1 = u_2$ on $[0, T - 2x_0 - 1]$. In terms of u this means $u(x) = u(x+1)$ for all $x \in [x_0, T - x_0 - 1]$. But that implies that

$$\int_x^{x+1} u = 0 \qquad \forall x \in [x_0, T - x_0 - 1].$$

It remains to be shown that

$$\int_x^{x+1} u > 0 \qquad \forall x \notin [x_0, T - x_0 - 1]. \tag{2.54}$$

By symmetry we only show this for $x < x_0$. Let u_p and g_p be the 1-periodic extrapolation of $u|_{[x_0, x_0+1]}$ and $g|_{[x_0, x_0+1]}$; note that $\int_x^{x+1} u_p = 0$ for every x. For $x < x_0$,

$$\begin{aligned} -2a(u)u'' - a'(u)u'^2 + b'(u) = 0 < g_p \\ = -2a(u_p)u_p'' - a'(u_p)u_p'^2 + b'(u_p), \end{aligned} \tag{2.55}$$

implying that $u > u_p$ for $x = x_0-$ and therefore also (2.54) for $x = x_0-$. If u and u_p intersect again at some $\tilde{x} < x_0$, then the comparison principle and (2.55) imply that $u \leq u_p$ on $[\tilde{x}, x_0]$, in contradiction with the previous statement. This concludes the proof. □

We discuss two different ways of calculating solutions of the problem (2.50–2.53).

Continuation

We implemented a strategy of continuation of solutions, using the continuation package AUTO [40], and we chose the simple case

$$a(u) = \frac{1}{2}, \qquad b(u) = \frac{1}{2}(u+1)^2. \tag{2.56}$$

To implement system (2.50–2.53) in AUTO, we divide $[0, T]$ into three subdomains, $[0, x_0]$, $[x_0, x_0 + 1]$ and $[x_0 + 1, T]$ and specify the equations

$$\left.\begin{aligned} -u_1''(x_1) + u_1(x_1) + 1 = \gamma(x_1; x_0, T, G) \\ x_1' = 1 \end{aligned}\right\} \text{ on } [0, x_0], \tag{2.57}$$

$$\left.\begin{aligned} -u_2''(x_2) + u_2(x_2) + 1 = \gamma(x_2; x_0, T, G) \\ x_2' = 1 \end{aligned}\right\} \text{ on } [x_0, x_0 + 1], \tag{2.58}$$

$$\left.\begin{aligned} -u_3''(x_3) + u_3(x_3) + 1 = \gamma(x_3; x_0, T, G), \\ x_3' = 1 \end{aligned}\right\} \text{ on } [x_0 + 1, T], \tag{2.59}$$

with boundary conditions

$$\begin{gathered} u_1(0) = 1, \\ u_1(x_0) = u_2(x_0),\ u_1'(x_0) = u_2'(x_0), \\ u_2(x_0) = u_3(x_0),\ u_2'(x_0) = u_3'(x_0), \\ u_3(T) = 1, \\ u_2(x_0) = u_2(x_0+1), \\ x_1(0) = 0,\ x_2(x_0) = x_0,\ x_3(x_0+1) = x_0+1. \end{gathered} \tag{2.60}$$

and integral condition

$$\int_{x_0}^{x_0+1} u_2 = 0. \tag{2.61}$$

Note that in (2.57)–(2.59) we have added trivial equations in order to solve for the x_i variables, which are required in the evaluation of $\bar{g}(x; x_0, G, T)$.

There are still some technicalities that have to be overcome: AUTO is not well-equipped to handle systems with a discontinuous right-hand side, such as the function $g(x; x_0, G, T)$ that is supplied here. We remedy this by using a low-order method for all simulations, and we smooth the function g given in (2.34) by substituting arctans for Heaviside functions:

$$\tilde{g}(x; x_0, T, G) =$$
$$\begin{cases} \frac{g_1}{\pi}\big(\arctan(A(x-x_0)) - \arctan(A(x-T-x_0))\big) & \text{if } T - 2x_0 \notin \mathbb{N}, \\ \quad + \frac{(g_2-g_1)}{\pi}\sum_{i=1}^{P}\big[\arctan(A(x-X_i)) - \arctan(A(x-Y_i))\big], & \\ \frac{1}{\pi}\big(\arctan(A(x-x_0)) - \arctan(A(x-T-x_0))\big) & \text{if } T - 2x_0 \in \mathbb{N}, \end{cases}$$

where X_i and Y_i are as in (2.29). In the limit $A \to \infty$, $\tilde{g}(x; x_0, T, G)$ converges pointwise to $\gamma(x; x_0, T, G)$.

There are nine differential equations with ten boundary conditions and one integral condition. This means that we expect to specify three free parameters to obtain a one-parameter curve of solutions. These are T, x_0, and an additional parameter β. It worked well to choose the freedom in β in modulating the values of $g_{1,2}$:

$$\tilde{g}_1 = g_1 + \beta \qquad \text{and} \qquad \tilde{g}_2 = g_2 + \beta.$$

One may prove *a priori* that $\beta = 0$ by noting that

$$\int_{x_0}^{x_0+1} \tilde{g} = \int_{x_0}^{x_0+1} \big[-u'' + u + 1\big] = 1,$$

and using (2.33) to find

$$1 = \int_{x_0}^{x_0+1} \tilde{g} = p(g_1 + \beta) + (1-p)(g_2 + \beta) = 1 + \beta.$$

We have found no other role for β than to accommodate for small numerical inaccuracies due to the discontinuous right-hand side. In all simulations $\beta \simeq 10^{-4}$.

We have validated the code by comparing solutions from AUTO with explicit solutions. An example is given in Figure 2.9.

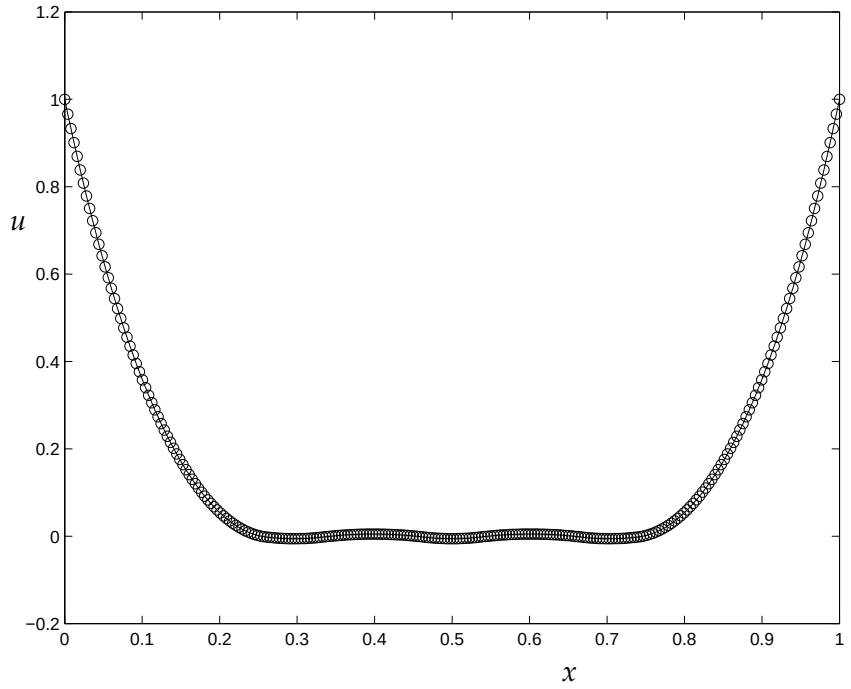


Figure 2.9: A comparison of a solution of system (2.57)–(2.61) produced with AUTO (○ symbols) to an explicit solution, for a generic value of T (here scaled to 1): $T = 4.91635$. In this simulation $A = 1000$.

As we have seen in the discussion at the beginning of this section, as T becomes larger the minimizer u has to have a contact point, and for large enough values even a full interval of contact. The point x_0, the leftmost point of contact, is determined as part of the solution; one may wonder how this point depends on T. For operators N that satisfy the comparison principle, it is straightforward to prove that x_0 remains

bounded for all T. Moreover, for the operator under consideration here, as $T \to \infty$, $x_0 \to \log(2+\sqrt{3})$. These two phenomena are illustrated in Figure 2.10.

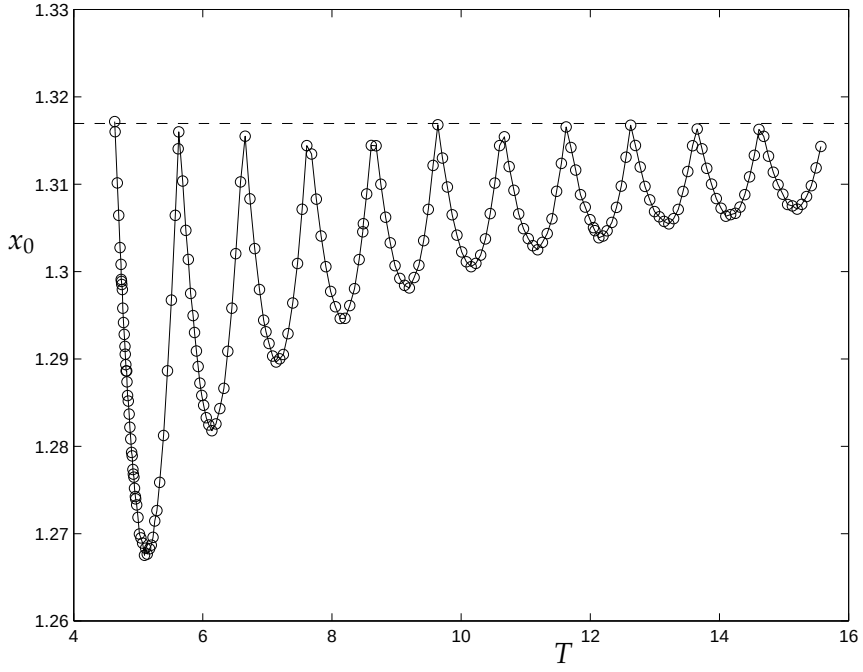


Figure 2.10: Behaviour of x_0 as a function of domain size T for system (2.57)–(2.61) computed with AUTO. As T grows, x_0 remains bounded and converges to $\log(2+\sqrt{3})$ (horizontal line).

Since $|g_1 - g_2| \to 0$ as P (and therefore T) increases, g becomes constant in the limit of large T. By the comparison principle, u does the same, implying that $F(u)/T \to 1$. The start of the convergence to 1 is shown in Figure 2.11.

Directly solving the boundary-value problem

Computing solutions of the rod equations—rather than the simpler problem (2.56)—using AUTO has proved difficult, for reasons that we do not understand well. Instead, a boundary-value problem solver from

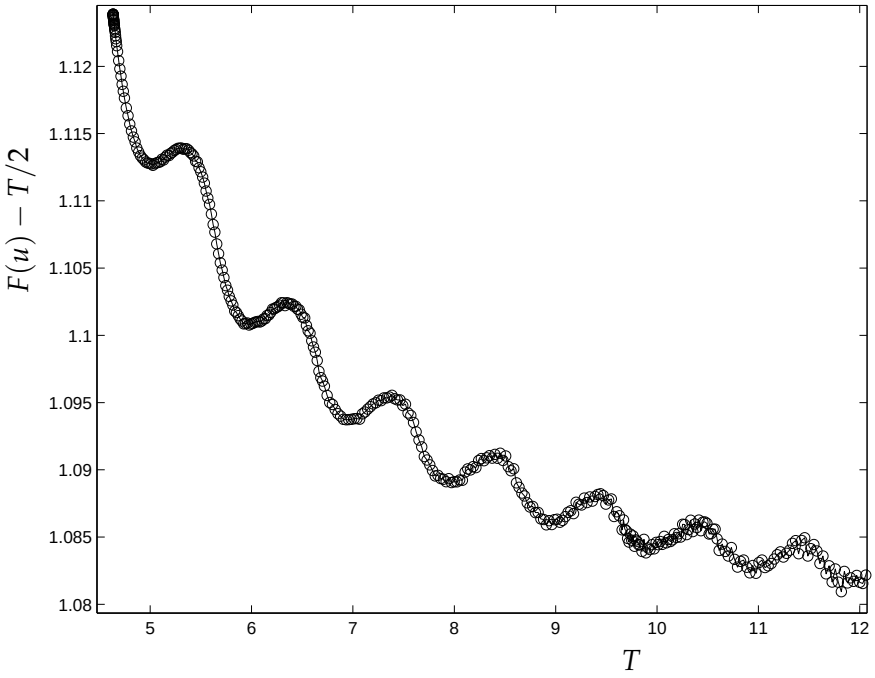


Figure 2.11: Behaviour of $F(u) - T/2$ as a function of domain size T for system (2.57)–(2.61) computed with AUTO. As T grows, $F(u) - T/2$ oscillates towards 1, the energy of $u \equiv 0$ on a unit length interval.

Matlab was used to create Figure 2.3. Set

$$Lu = -\frac{2u''}{4\pi^2(1+u^2)^{\frac{5}{2}}} + \frac{5uu'^2}{4\pi^2(1+u^2)^{\frac{7}{2}}} - \frac{3u}{r^2(1+u^2)^{\frac{5}{2}}} + \frac{\alpha}{(1+u^2)^{\frac{3}{2}}}.$$

To find a solution of

$$Lu = g(x; x_0, T, G), \quad u(x_0) = u(x_0+1), \quad \int_{x_0}^{x_0+1} u = 0,$$

for a generic value of T (large enough) we construct a two-parameter shooting problem. Fix G and x_0 and consider the boundary-value problem

$$\left.\begin{array}{ll} Lu_1 = 0 & \text{on } [0, x_0], \\ Lu_2 = g_1 & \text{on } [x_0, x_0+p], \\ Lu_3 = g_2 & \text{on } [x_0+p, x_0+1], \\ Lu_u = g_1 & \text{on } [x_0+1, x_0+1+p], \\ Lu_5 = 0 & \text{on } [x_0+1+p, \tilde{T}], \end{array}\right\} \tag{2.62}$$

with boundary conditions

$$u_1(0) = 1, \tag{2.63}$$

$$u_1(x_0) = u_2(x_0),\ u_1'(x_0) = u_2'(x_0), \tag{2.64}$$

$$u_2(x_0 + p) = u_3(x_0 + p),\ u_2'(x_0 + p) = u_3'(x_0 + p), \tag{2.65}$$

$$u_3(x_0 + 1) = u_4(x_0 + 1),\ u_3'(x_0 + 1) = u_4'(x_0 + 1), \tag{2.66}$$

$$u_4(x_0 + 1 + p) = u_5(x_0 + 1 + p),\ u_4'(x_0 + 1 + p) = u_5'(x_0 + 1 + p), \tag{2.67}$$

$$u_5(\tilde{T}) = 1. \tag{2.68}$$

Here, as before, $p \equiv T - 2x_0 - 1 \pmod 1$, $P = \min\{n \in \mathbb{N} : n \geq T - 2x_0 - 1\}$, and

$$g_1 = \frac{GP}{P + 1 - p}, \quad g_2 = \frac{G(P+1)}{P + 1 - p},$$

by Theorem 2.17. Note that this is not exactly the same problem as (2.50–2.51), since the periodic section has been reduced from P periods to a single period, and the solution is defined correspondingly on a smaller domain of length

$$\tilde{T} = 2x_0 + p + 1.$$

This allows us to use the decomposition in five subdomains for any T, which facilitates computation. This is illustrated in Figure 2.12.

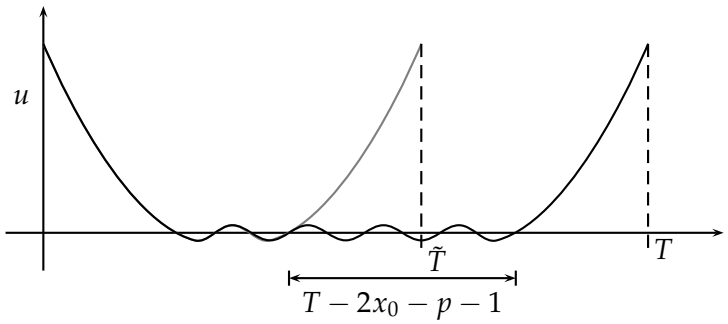


Figure 2.12: Schematic picture of the idea behind $\tilde{T} = 2x_0 + p + 1$. Since u (solid black line) is periodic between x_0 and $T - x_0$, we can cut out an interval of length $T - 2x_0 - p - 1$ and find the corresponding solution on $[0, \tilde{T}]$.

We now vary x_0 and G to find solutions of system (2.62)–(2.68) that satisfy

$$u_2(x_0) = u_3(x_0 + 1), \quad \int_{x_0}^{x_0+p} u_2 + \int_{x_0+p}^{x_0+1} u_3 = 0,$$

using a standard Matlab boundary-value problem solver, `bvp4c`. An example solution is drawn in Figure 2.3 in which we have used $\alpha = 1/2\pi$, $r = 1$. Note that all analysis in this study assumes zero rod thickness; in Figure 2.3 the rod has been artificially fattened for better viewing.

CHAPTER 3

Link, twist, and writhe for open rods

3.1. Introduction

In a variational analysis of an elastic structure that is acted upon by end forces and moments one needs to consider the work done by the applied loads. The work done by the applied force does not usually present any problems. One requires the distance travelled by the force, which is usually easy to obtain. More problems occur in determining the work done by the applied moment. Here one requires the end rotation as 'seen' by the moment, which may lead to complications if large deformations are allowed. This paper discusses the ambiguities associated with this end rotation and shows how a consistent treatment can be obtained.

If the structure is very long it may be modelled by an infinitely long rod. This has the advantage that powerful techniques from dynamical systems (viewing arclength along the rod as time) [26, 70] and variational analysis [108] can be used. The natural class of solutions to consider in this case are localised solutions, which decay sufficiently rapidly towards the ends (other solutions would have infinite strain energy and would therefore be non-physical). The boundary conditions to be imposed on such solutions are simple: end tangents are aligned and the ends of the rod do not interfere with the *localised* deformation. The work done by the torsional load is then simply the product of applied end twisting moment M and relative end rotation R (the angle that has gone into one end of the rod in order to produce the deformation, starting from a straight and untwisted reference configuration and keeping the other end fixed).

This chapter is based on [67].

However, any real-world problem (be it a drill string, a marine pipeline or a DNA supercoil) deals with a finite-length rod. For such a rod more complicated boundary conditions may be encountered. For instance, the end tangents need not be aligned, so that an end rotation that could be used in an energy discussion is not straightforward to define. But even in the case of aligned end tangents complications arise if large deformations are allowed. These complications are the subject of this paper.

Let us demonstrate the issue with an example. If we rotate the right end of the rod in Fig. 3.1(a) through an angle of -4π we obtain the configuration shown in (b). If we now move the right end of the rod to the left, the rod pops into a looped configuration as shown in (c). If we move the right end further and make the loop pass around the right clamp, as in (d), and then pull the ends out again, we return to configuration (a) without having rotated the ends. (Note that the process illustrated in Fig. 3.1(d) requires that whatever supports the right clamp must be released to allow the passage of the rod behind it.) We conclude that we can go from configuration (a) to configuration (b) either by end rotation or by 'looping'. What, then, is the 'real' end rotation of the deformation (a) $\rightarrow$ (b) 'seen' by the applied moment and therefore pertinent to an energy analysis?

It is apparent from this ambiguity that it is not sufficient, *a priori*, to consider only the initial and final configurations of a deformation process to decide on the end rotation. We need to be told *how* the rod was deformed from the one into the other, i.e., we need to know the deformation *history*, not just locally of the ends but globally of the entire rod. (This path dependence suggests a relation with a geometric phase; details on this are found in [62, 93, 128].)

Alexander & Antman [4] address the end rotation ambiguity for fixed boundary conditions by imagining the open rod to be part of a closed rod. A natural restriction to a class of open rod deformations is then obtained by demanding that the closed continuation should not undergo self-intersections. Applied to Fig. 3.1 this would mean that throwing the rod around the clamp as in (d) is excluded because it requires an intersection with the closure (indicated by the dotted lines). In the restricted class of deformations the configurations (b) and (c) can then be assigned a unique end rotation of $R = -4\pi$ relative to (a). The ambiguity in the end rotation is thus resolved.

To distinguish between different classes of deformations Alexander & Antman use the *link* of the closed rod, i.e., the topological linking

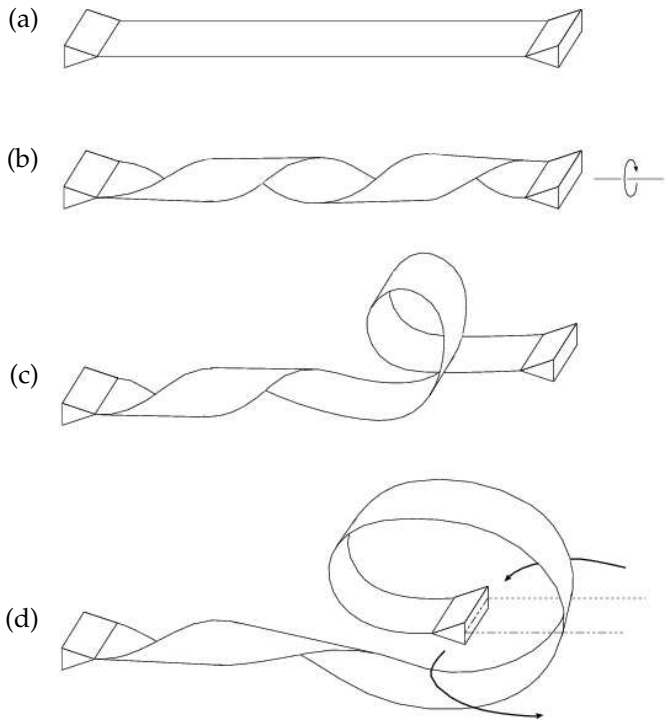


Figure 3.1: Four rod configurations that are not distinguished by a prescribed process (homotopy) of boundary conditions from a fixed reference configuration: one can go from (a) to (b) either through an end rotation or through 'looping'. (After [4]).

number of two lines drawn on opposite sides of the unstressed rod. Since two unknotted closed rods with equal linking number may be deformed into each other without undergoing self-intersections, there is a one-to-one correspondence between the classes of admissible deformations with unique R and the values of the link. The link is thus seen to be related to the end rotation. (Note that the -4π versus 0 ambiguity in the end rotation illustrated in Fig. 3.1 is reflected in the jump in link by -2 as the closed rod intersects itself.)

The connection between link for closed rods and end rotation for open rods has been discussed by many authors. The level of detail varies, but one usually argues from the link of a closed rod to the link of an open rod, which is then identified with the end rotation. Often this involves the introduction of a closure and the use of the celebrated

formula [22, 50, 142]

$$Lk = Tw + Wr, \tag{3.1}$$

which expresses the link Lk (of a closed rod) in terms of the twist Tw (a local property, in the sense that it can be found by integrating a density along the length of the rod) and the writhe Wr. This writhe, which is only a function of the centreline of the rod, is not a local property but several expressions exist for the calculation of the writhe of an arbitrary closed curve [3]. Some of these expressions make sense for open curves as well and therefore suggest an extension of writhe from closed to open curves [129]. The final result, then, is a formula for the open link which in terms of a suitable Euler-angle representation takes the simple form

$$\text{open link} = \frac{1}{2\pi}\int[\dot{\phi} + \dot{\psi}]. \tag{3.2}$$

This formula cannot be expected to hold true in general since it would make link a local property, which it is not. Indeed, the generalisation of writhe from closed to open curves used in obtaining (3.2) is subject to a geometrical condition (see Section 3.2). Certainly general validity of (3.2) is prevented by the polar singularity inevitably associated with Euler angles. In Section 3.3 formula (3.2) will be derived within a limited class of deformations.

The present study improves on previous results in the following ways.

(1) The closure introduced by Alexander & Antman is effective in the definition of admissible deformations if the supports are fixed in both angle and position. This is adequate for a large class of experimental situations. If, however, the supports are allowed to move, then it is not possible to use one and the same closure for all deformations: different deformations require different closures. Thus the desired distinction between classes of deformations is lost (one can always move the closure along, so that it does not 'get in the way' during the deformation). We resolve the issue by limiting the class of admissible closures in such a way that the separation into different classes of deformations, characterised by the link, is preserved under arbitrary movement of the supports.

(2) We define a precise class of deformations within which link, twist, and writhe are well-defined. We carefully examine the restrictions imposed on the deformations and show in which sense they are necessary. As mentioned above, for a consistent definition of link and writhe it is necessary to work within a class of *homotopies of*

rods, which connect a given open rod to a reference configuration. Despite the requirement of such a connection, the newly introduced writhe and link themselves are independent of the choice of connection: within the class the writhe and link only depend on the given open rod and the reference configuration.

(3) We show that the link, twist and writhe defined for open rods satisfy the classical equality (3.1). Furthermore, within the class of admissible deformations the link is given by (3.2). In the special case that the end tangents of the open rod are aligned the link coincides with the end rotation. Our results thus formalise current practise in the literature based on (3.2). However, while most applications of formula (3.2) can be shown to fall into our class of deformations, recent experiments in molecular biology do not always do so and care is required in energy discussions (we discuss this in Section 4). Indeed, we would claim that usage of (3.2) presupposes the framework that we here discuss, and consequently is subject to the limitations that we describe.

The organisation of this chapter is as follows. In Section 3.2 we first define our class of open rods. For the elements of this class we define link, twist and writhe and derive the extension of (3.1) to open rods. In Section 3.3 we independently define the end rotation and show it to be equal to the open link. We also derive (3.2). In Section 3.4 we critically review the defining conditions of the class of rods considered, illustrating their relevance with counterexamples. Section 3.5 discusses our work in the light of previous work in the literature.

3.2. Results: Link, Twist and Writhe

For our purposes a rod is a member of the set

$$\mathcal{A}^0 = \{\, (\boldsymbol{r}, \boldsymbol{d}_1) \in C^2([0,L];\mathbb{R}^3 \times S^2) \text{ such that } |\dot{\boldsymbol{r}}| \neq 0,\ \dot{\boldsymbol{r}} \cdot \boldsymbol{d}_1 = 0, \\ \text{and } \boldsymbol{r} \text{ is non-self-intersecting} \,\}.$$

Here and in the following an overdot denotes differentiation with respect to the spatial variable s. The curve $\boldsymbol{r}$ is thought of as the centreline of a physical rod (of length L) and $\boldsymbol{d}_1(s)$ as a material vector in the section at s. As alternatives to 'rod' the terms 'ribbon' [51, 3] and 'strip' [4] are also used. A closed rod is an element of $\mathcal{A}^0$ for which begin and end connect smoothly.

To each point on the centreline of the rod we can attach an orthonormal right-handed frame $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ of directors by setting

$$\boldsymbol{d}_3(s) = \dot{\boldsymbol{r}}(s)/|\dot{\boldsymbol{r}}(s)| \qquad \text{and} \qquad \boldsymbol{d}_2(s) = \boldsymbol{d}_3(s) \times \boldsymbol{d}_1(s). \tag{3.3}$$

These directors track the varying orientation of the cross-section of the rod along the length of the rod. The twist of a closed rod $(\boldsymbol{r}, \boldsymbol{d}_1)$ is now defined by

$$Tw(\boldsymbol{r}, \boldsymbol{d}_1) := \frac{1}{2\pi} \oint_{\boldsymbol{r}} \dot{d}_1(s) \cdot \boldsymbol{d}_2(s)\, ds. \tag{3.4}$$

It measures the number of times $\boldsymbol{d}_1$ revolves around $\boldsymbol{d}_3$ in the direction of $\boldsymbol{d}_2$ as we go around the rod.

Let $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ be two non-intersecting curves. Then the link of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ is defined by

$$Lk(\boldsymbol{r}_1, \boldsymbol{r}_2) := \frac{1}{4\pi} \oint_{\boldsymbol{r}_1} \oint_{\boldsymbol{r}_2} \frac{[\dot{\boldsymbol{r}}_1(s) \times \dot{\boldsymbol{r}}_2(t)] \cdot [\boldsymbol{r}_1(s) - \boldsymbol{r}_2(t)]}{|\boldsymbol{r}_1(s) - \boldsymbol{r}_2(t)|^3}\, dsdt. \tag{3.5}$$

The writhe of a closed curve $\boldsymbol{r}$ is

$$Wr(\boldsymbol{r}) := \frac{1}{4\pi} \oint_{\boldsymbol{r}} \oint_{\boldsymbol{r}} \frac{[\dot{\boldsymbol{r}}(s) \times \dot{\boldsymbol{r}}(t)] \cdot [\boldsymbol{r}(s) - \boldsymbol{r}(t)]}{|\boldsymbol{r}(s) - \boldsymbol{r}(t)|^3}\, dsdt. \tag{3.6}$$

The argument of this integral is the pullback of the area form on S^2 under the Gauss map,

$$\begin{array}{rccc} G: & \mathbb{R}^2 & \longrightarrow & S^2 \\ & (\boldsymbol{r}(s), \boldsymbol{r}(t)) & \longmapsto & \frac{\boldsymbol{r}(s)-\boldsymbol{r}(t)}{|\boldsymbol{r}(s)-\boldsymbol{r}(t)|}, \end{array}$$

so that the writhe may be interpreted as the signed area on S^2 that is covered by this map. For each direction $p \in S^2$ the signed multiplicity of the Gauss map (i.e., the number of points (s,t) for which $G(s,t) = p$, weighted by the sign of $p \cdot [G_s \times G_t]$) equals the *directional writhing number*, the number of signed crossings of the projection of $\boldsymbol{r}$ onto a plane orthogonal to the vector p [50, 3]. In other words, the writhe of a closed curve is equal to the directional writhing number averaged over all directions of S^2.

The link, twist and writhe of a closed rod are related by the well-known Călugăreanu-White-Fuller Theorem [22, 142, 50]:

THEOREM 3.1. *Let $(\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^0$ be a closed rod as defined above. Then*

$$Lk(\boldsymbol{r}, \boldsymbol{d}_1) = Tw(\boldsymbol{r}, \boldsymbol{d}_1) + Wr(\boldsymbol{r}). \tag{3.7}$$

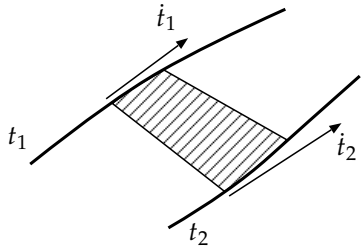


Figure 3.2: The argument of the integral in (3.9) is the area swept out by the geodesic connecting the curves $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ on S^2.

We review two classical theorems by Fuller which are of interest to us. Note that at each point $\boldsymbol{r}(s)$, the unit tangent $t(s) = \dot{\boldsymbol{r}}(s)/|\dot{\boldsymbol{r}}(s)|$ traces out a closed curve on S^2, called the tantrix. Fuller's first theorem relates the writhe of the curve to the area A enclosed by the tantrix on S^2:

$$Wr(\boldsymbol{r}) = \frac{A}{2\pi} - 1 \ (\text{modd } 2). \tag{3.8}$$

Note that the equality modulo two is necessary since the area enclosed by a curve on S^2 is only defined modulo 4π.

The second theorem, stated in detail as Theorem 3.6 below, gives under certain conditions a formula for the difference in writhe between two closed curves $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ that can be continuously deformed into each another (see Figure 3.2):

$$Wr(\boldsymbol{r}_1) - Wr(\boldsymbol{r}_2) = \frac{1}{2\pi} \int \frac{\boldsymbol{t}_2 \times \boldsymbol{t}_1}{1 + \boldsymbol{t}_1 \cdot \boldsymbol{t}_2} \cdot (\dot{\boldsymbol{t}}_1 + \dot{\boldsymbol{t}}_2). \tag{3.9}$$

We now proceed with the introduction of the set for which the open link, twist and writhe will be defined. For the definition of this set we choose a closed planar reference curve $\boldsymbol{r}_0 \in C^2([0, M]; \mathbb{R}^3)$ for some $M > L$.

DEFINITION 3.2.

$$\mathcal{A}^1_{\boldsymbol{r}_0} = \Big\{ \quad (\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^0 \text{ such that } \exists (\bar{\boldsymbol{r}}, \bar{\boldsymbol{d}}_1) \in C^2([0,M]\times[0,1]; \mathbb{R}^3 \times S^2) :$$

1. *for each* λ, $\bar{\boldsymbol{r}}(\cdot, \lambda)$ *is an unknotted, non-self-intersecting closed curve,*
2. $\dot{\bar{\boldsymbol{r}}}(s,\lambda) \cdot \bar{\boldsymbol{d}}_1(s,\lambda) = 0 \quad$ *for* $s \in [0,M]$, $\lambda \in \{0,1\}$,
3. $(\bar{\boldsymbol{r}}, \bar{\boldsymbol{d}}_1)(s,1) = (\boldsymbol{r}, \boldsymbol{d}_1)(s) \quad$ *for* $s \in [0,L]$,
4. $\bar{\boldsymbol{r}}(s,0) = \boldsymbol{r}_0(s) \quad$ *for* $s \in [0,L]$,
5. $\dot{\bar{\boldsymbol{r}}}(s,0) \cdot \dot{\bar{\boldsymbol{r}}}(s,\lambda) > -1 \quad$ *for* $s \in [0,M]$, $\lambda \in [0,1]$,
6. $\{\bar{\boldsymbol{r}}(s,\lambda) \colon s \in [L,M]\}$ *is a planar curve for* $\lambda = 0$ *and* $\lambda = 1$, *and these two planes are parallel* $\Big\}$.

$\mathcal{A}^1_{\boldsymbol{r}_0}$ can be thought of as a class of open rods $(\boldsymbol{r}, \boldsymbol{d}_1)$ that can be connected by a homotopy—satisfying certain requirements—to the reference curve $\boldsymbol{r}_0$. The part of the closed rod parametrised by $s \in [L, M]$ is called the closure.

Some of the conditions above are more straightforward than others. Parts 1 and 2 state that $(\bar{\boldsymbol{r}}, \bar{\boldsymbol{d}}_1)$ is a homotopy of well-behaved closed rods, and by parts 3 and 4 the homotopy contains the original open rod $(\boldsymbol{r}, \boldsymbol{d}_1)$ at $\lambda = 1$, and the reference curve at $\lambda = 0$. Parts 5 and 6 contain the essential elements of this definition. Part 5 is the same non-opposition condition that appears in the statement of Fuller's theorem (Theorem 3.6) and is required for the conversion of the writhe to a single-integral expression. Part 6, which states that the closure should be planar at the beginning and the end of the homotopy, is central in the construction. These last two conditions are discussed more fully in Section 3.4.

Note that curves $\boldsymbol{r}_0$ exist for which $\mathcal{A}^1_{\boldsymbol{r}_0}$ is empty: if the three vectors $\dot{\boldsymbol{r}}_0(0)$, $\dot{\boldsymbol{r}}_0(L)$, and $\boldsymbol{r}_0(L) - \boldsymbol{r}_0(0)$ are independent, then the open curve $\boldsymbol{r}_0$ can not be closed by a planar closure, so that the set of homotopies with planar closures that connect to $\boldsymbol{r}_0$ is empty.

We are now in a position to define the new functionals open link, open twist and open writhe for open rods.

DEFINITION 3.3. *Let $(\boldsymbol{r}, \boldsymbol{d}_1)$ be a rod in $\mathcal{A}^1_{\boldsymbol{r}_0}$. Then the open twist of $(\boldsymbol{r}, \boldsymbol{d}_1)$ is*

$$Tw^o(\boldsymbol{r}, \boldsymbol{d}_1) := \frac{1}{2\pi} \int_0^L \dot{\boldsymbol{d}}_1 \cdot (\dot{\boldsymbol{r}} \times \boldsymbol{d}_1)\, ds,$$

the open writhe of $\boldsymbol{r}$ is

$$Wr^o(\boldsymbol{r}) := Wr(\bar{\boldsymbol{r}}(\cdot, 1)) = \frac{1}{4\pi} \int_0^M \int_0^M \frac{[\bar{\boldsymbol{r}}(s,1) - \bar{\boldsymbol{r}}(t,1)] \cdot [\dot{\bar{\boldsymbol{r}}}(s,1) \times \dot{\bar{\boldsymbol{r}}}(t,1)]}{|\bar{\boldsymbol{r}}(s,1) - \bar{\boldsymbol{r}}(t,1)|^3}\, dsdt,$$

and the open link of $(\boldsymbol{r}, \boldsymbol{d}_1)$ is

$$Lk^o(\boldsymbol{r}, \boldsymbol{d}_1) := Lk(\bar{\boldsymbol{r}}(\cdot, 1), \bar{\boldsymbol{d}}_1(\cdot, 1)) - \frac{1}{2\pi} \int_L^M \dot{\bar{\boldsymbol{d}}}_1(s,1) \cdot (\dot{\bar{\boldsymbol{r}}}(s,1) \times \dot{\bar{\boldsymbol{d}}}_1(s,1))\, ds.$$

Note that in the last definition we subtract any twist the closure might have. It follows directly from the construction that the new concepts also satisfy the classical relationship:

COROLLARY 3.4. *Let $(\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^1_{\boldsymbol{r}_0}$ be an open rod. Then*

$$Lk^o(\boldsymbol{r}, \boldsymbol{d}_1) = Tw^o(\boldsymbol{r}, \boldsymbol{d}_1) + Wr^o(\boldsymbol{r}).$$

THEOREM 3.5. *For any open rod $(\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^1_{\boldsymbol{r}_0}$, the open twist, writhe, and link are well-defined.*

For writhe and link this is non-trivial, as different homotopy closures $(\bar{\boldsymbol{r}}, \bar{\boldsymbol{d}}_1)$ might be expected to give rise to different values.

PROOF. We first state Fuller's second theorem in a more precise form.

THEOREM 3.6. *Let $\boldsymbol{r}_\lambda$ $(0 \le \lambda \le 1)$ be a homotopy of closed non-self-intersecting curves, regularly parametrized with a common parameter $s \in [0, L]$. Let $\boldsymbol{t}_\lambda$ be the tantrix of $\boldsymbol{r}_\lambda$. If $\boldsymbol{t}_0(s) \cdot \boldsymbol{t}_\lambda(s) > -1$ for all $s \in [0, L]$, $\lambda \in [0, 1]$, then*

$$Wr(\boldsymbol{r}_1) - Wr(\boldsymbol{r}_0) = \frac{1}{2\pi} \int_0^L \frac{\boldsymbol{t}_0 \times \boldsymbol{t}_1}{1 + \boldsymbol{t}_0 \cdot \boldsymbol{t}_1} \cdot (\dot{\boldsymbol{t}}_0 + \dot{\boldsymbol{t}}_1) ds. \tag{3.10}$$

To our knowledge, a rigorous proof was first given by Aldinger *et al.* [3].

To prove Theorem 3.5 for the open writhe, let $(\bar{\boldsymbol{r}}, \bar{\boldsymbol{d}}_1)$ be a homotopy associated to $(\boldsymbol{r}, \boldsymbol{d}_1)$. By definition, $\bar{\boldsymbol{r}}(\cdot, 0)$ and $\bar{\boldsymbol{r}}(\cdot, 1)$ are planar for $s \in [L, M]$. Let us denote the planes by V_0 and V_1; these are parallel by Definition 3.2.6. Let V be the plane through the origin parallel to both.

We have defined the class of open rods $\mathcal{A}^1_{r_0}$ such that Theorem 3.6 can be applied. Denote the tantrices of $\bar{\boldsymbol{r}}(\cdot, 0)$ and $\bar{\boldsymbol{r}}(\cdot, 1)$ by $\boldsymbol{t}_0$ and $\boldsymbol{t}_1$ respectively. Then

$$Wr(\bar{\boldsymbol{r}}(\cdot, 1)) - Wr(\bar{\boldsymbol{r}}(\cdot, 0)) = \frac{1}{2\pi} \int_0^M \frac{\boldsymbol{t}_0(s) \times \boldsymbol{t}_1(s)}{1 + \boldsymbol{t}_0(s) \cdot \boldsymbol{t}_1(s)} \cdot (\dot{\boldsymbol{t}}_0(s) + \dot{\boldsymbol{t}}_1(s))\, ds. \quad (3.11)$$

The argument of the integral vanishes for $s \in [L, M]$: since $\boldsymbol{t}_0(s), \boldsymbol{t}_1(s) \in V$ for $s \in [L, M]$ we have $\boldsymbol{t}_0(s) \times \boldsymbol{t}_1(s) \perp V$ and $\dot{\boldsymbol{t}}_0(s) + \dot{\boldsymbol{t}}_1(s) \in V$ for $s \in [L, M]$. Hence

$$[\boldsymbol{t}_0(s) \times \boldsymbol{t}_1(s)] \cdot [\dot{\boldsymbol{t}}_0(s) + \dot{\boldsymbol{t}}_1(s)] = 0.$$

Moreover, since $\bar{\boldsymbol{r}}(\cdot, 0)$ is planar, $Wr(\bar{\boldsymbol{r}}(\cdot, 0)) = 0$. We conclude

$$Wr(\bar{\boldsymbol{r}}(\cdot, 1)) = \frac{1}{2\pi} \int_0^L \frac{\boldsymbol{t}_0(s) \times \boldsymbol{t}_1(s)}{1 + \boldsymbol{t}_0(s) \cdot \boldsymbol{t}_1(s)} \cdot (\dot{\boldsymbol{t}}_0(s) + \dot{\boldsymbol{t}}_1(s))\, ds. \quad (3.12)$$

Since this integral only depends on the reference curve and the open rod itself, and is otherwise independent of the choice of closure and homotopy, this proves the claim for the writhe.

For the link, let $(\bar{\boldsymbol{r}}, \bar{\boldsymbol{d}}_1)$ be a homotopy associated to $(\boldsymbol{r}, \boldsymbol{d}_1)$ by Definition 3.2. Denote the closed curves $\bar{\boldsymbol{r}}(\cdot, 1)$ and $\bar{\boldsymbol{d}}_1(\cdot, 1)$ by $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{d}}_1$ respectively. We denote the open twist evaluated over an interval $s \in [a, b]$ by $Tw^o_{[a,b]}$. Then

$$\begin{aligned} Tw^o_{[L,M]}(\hat{\boldsymbol{r}}, \hat{\boldsymbol{d}}_1) &= Tw^o_{[0,M]}(\hat{\boldsymbol{r}}, \hat{\boldsymbol{d}}_1) - Tw^o_{[0,L]}(\hat{\boldsymbol{r}}, \hat{\boldsymbol{d}}_1) \\ &= Tw^o_{[0,M]}(\hat{\boldsymbol{r}}, \hat{\boldsymbol{d}}_1) - Tw^o(\boldsymbol{r}, \boldsymbol{d}_1). \end{aligned}$$

Hence

$$\begin{aligned} Lk^o(\boldsymbol{r}, \boldsymbol{d}_1) &= Lk(\hat{\boldsymbol{r}}, \hat{\boldsymbol{d}}_1) - Tw^o_{[L,M]}(\hat{\boldsymbol{r}}, \hat{\boldsymbol{d}}_1) \\ &= Lk(\hat{\boldsymbol{r}}, \hat{\boldsymbol{d}}_1) - Tw(\hat{\boldsymbol{r}}, \hat{\boldsymbol{d}}_1) + Tw^o(\boldsymbol{r}, \boldsymbol{d}_1) \\ &= Wr(\hat{\boldsymbol{r}}) + Tw^o(\boldsymbol{r}, \boldsymbol{d}_1) \\ &= Wr^o(\boldsymbol{r}) + Tw^o(\boldsymbol{r}, \boldsymbol{d}_1). \end{aligned} \quad (3.13)$$

It follows that $Lk^o(\boldsymbol{r}, \boldsymbol{d}_1)$ is independent of the chosen closure $(\bar{r}, \bar{\boldsymbol{d}}_1)$. □

Equation (3.12) is an important motivation of this work, since it expresses the writhe in terms of a single rather than a double integral. For the purpose of variational analysis this is an obvious advantage. It is especially useful when the link, and therefore indirectly the writhe, can be identified with the rotation of the ends; this requires that the end tangents remain equal throughout the deformation, and this case is treated in the next section.

REMARK 3.7. For simplicity we have chosen to introduce one class $\mathcal{A}^1_{\boldsymbol{r}_0}$ as the basis for the definitions of open link, twist and writhe. For each of the three definitions separately, however, not all of Definition 3.2 is required. Open twist can be defined directly in terms of $(\boldsymbol{r}, \boldsymbol{d}_1)$, without the need for an extension; open link requires a closure, but no homotopy; and open writhe requires a homotopy, but the director $\boldsymbol{d}_1$ can be disposed of.

REMARK 3.8. A natural question to ask is whether the open link, twist and writhe reduce to their classical counterparts when an open rod is transformed into a closed rod by lining up and connecting the ends. This is not the case, as we demonstrate in Section 3.4.

3.3. Results: End-rotation and Euler angles

It is common in applications to assume that the end tangents of the buckled rod are kept constant and equal during the deformation process. For comparison with an end rotation we introduce this additional condition. Throughout this section we also assume that $\boldsymbol{r}_0|_{[0,L]}$ is straight; without loss of generality we assume that $\dot{\boldsymbol{r}}_0$ is a constant unit vector v on $[0, L]$. Finally, again without loss of generality we choose the director $\boldsymbol{d}_1$ constant on the reference curve $\boldsymbol{r}_0|_{[0,L]}$:

DEFINITION 3.9.

$$\begin{aligned}\mathcal{A}^2_{\boldsymbol{r}_0} = \{\, (\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^1_{\boldsymbol{r}_0} :\ & \dot{\boldsymbol{r}}_0(s) = v \in S^2 \textit{ for all } s \in [0, L], \\ & \dot{\bar{\boldsymbol{r}}}(0, \lambda) = \dot{\bar{\boldsymbol{r}}}(L, \lambda) = \dot{\bar{\boldsymbol{r}}}(M, \lambda) = v \textit{ for all } \lambda \in [0, 1], \\ & \bar{\boldsymbol{d}}_1(s, 0) = \bar{\boldsymbol{d}}_1(0, 0) \textit{ for all } s \in [0, L] \,\}.\end{aligned}$$

The following formula is a direct consequence of Theorem 3.6:

COROLLARY 3.10. *Let* $(\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^2_{\boldsymbol{r}_0}$ *and let* $\boldsymbol{t}$ *be the tantrix of* $\boldsymbol{r}$*. Then*

$$Wr^o(\boldsymbol{r}) = \frac{1}{2\pi} \int_0^L \frac{v \times \boldsymbol{t}(s)}{1 + v \cdot \boldsymbol{t}(s)} \cdot \dot{\boldsymbol{t}}(s)\, ds. \tag{3.14}$$

In the present case of a straight $\boldsymbol{r}_0|_{[0,L]}$ the dependence on $\boldsymbol{r}_0$ of the open writhe of a given open rod takes a particularly simple form:

THEOREM 3.11. *Under the conditions of Corollary 3.10, let* $\Omega = S^2 \setminus \{-\boldsymbol{t}(s) : s \in [0, L]\}$*. Then the function*

$$v \in S^2 \longmapsto \frac{1}{2\pi} \int_0^L \frac{v \times \boldsymbol{t}(s)}{1 + v \cdot \boldsymbol{t}(s)} \cdot \dot{\boldsymbol{t}}$$

(s) ds

is constant on connected components of Ω*.*

The proof is given in the appendix. The interpretation of this theorem is as follows: when the end tangents are aligned, the tantrix given by the rod (without closure) forms a closed curve on S^2. The integral above represents 'area enclosed by the curve' for a given 'choice of area' (*cf.* (3.8)). When the vector v crosses the set $\{-\boldsymbol{t}(s) : s \in [0, L]\}$ the geodesic connections between v and $\boldsymbol{t}(s)$ change direction, causing the integral to represent a different choice of area, and therefore causing the integral to jump by 4π.

With fixed end tangents we can introduce a fourth quantity, the end rotation. We denote $\partial(\cdot)/\partial\lambda$ by $\partial_\lambda(\cdot)$.

DEFINITION 3.12. *Let* $(\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^2_{\boldsymbol{r}_0}$*, and let* $\bar{\boldsymbol{d}}_3(\cdot,\cdot) = \dot{\bar{\boldsymbol{r}}}(\cdot,\cdot)/\,|\dot{\bar{\boldsymbol{r}}}(\cdot,\cdot)|$*,* $\bar{\boldsymbol{d}}_2 = \bar{\boldsymbol{d}}_3 \times \bar{\boldsymbol{d}}_1$*. We define the end rotation by*

$$R(\boldsymbol{r}, \boldsymbol{d}_1) := \int_0^1 \partial_\lambda \bar{\boldsymbol{d}}_1(L, \lambda) \cdot \bar{\boldsymbol{d}}_2(L, \lambda)\, d\lambda - \int_0^1 \partial_\lambda \bar{\boldsymbol{d}}_1(0, \lambda) \cdot \bar{\boldsymbol{d}}_2(0, \lambda)\, d\lambda.$$

To study the relationship between end rotation and open link, twist and writhe we introduce a particular choice of Euler angles for an open rod $(\boldsymbol{r}, \boldsymbol{d}_1)$. Recall that for every $s \in [0, L]$ there is an orthonormal director frame $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$. We express this frame in terms of angles

θ, ψ, ϕ with respect to a fixed basis $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ as follows

$$\begin{aligned}
\boldsymbol{d}_1 &= (-\sin\psi\sin\phi + \cos\psi\cos\phi\cos\theta)\,\boldsymbol{e}_1 + \\
&\quad (\cos\psi\sin\phi + \sin\psi\cos\phi\cos\theta)\,\boldsymbol{e}_2 - \cos\phi\sin\theta\,\boldsymbol{e}_3, \\
\boldsymbol{d}_2 &= (-\sin\psi\sin\phi - \cos\psi\sin\phi\cos\theta)\,\boldsymbol{e}_1 + \\
&\quad (\cos\psi\cos\phi + \sin\psi\sin\phi\cos\theta)\,\boldsymbol{e}_2 - \sin\phi\sin\theta\,\boldsymbol{e}_3, \\
\boldsymbol{d}_3 &= \cos\psi\sin\theta\,\boldsymbol{e}_1 + \sin\psi\sin\theta\,\boldsymbol{e}_2 + \cos\theta\,\boldsymbol{e}_3.
\end{aligned} \tag{3.15}$$

This choice of Euler angles follows Love [90, art. 253]. For rods in the class $\mathcal{A}^2_{r_0}$ we choose $\boldsymbol{e}_3$ parallel to v; note that by this choice the non-opposition condition 5 in Definition 3.2 coincides with avoidance of the Euler-angle singularity at $\theta = \pi$. Therefore the smoothness assumptions on $(\bar{\boldsymbol{r}}, \bar{\boldsymbol{d}}_1)$ in $\mathcal{A}^2_{r_0}$ imply C^1-regularity for ϕ, ψ, and θ.

LEMMA 3.13. *Let $(\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^2_{r_0}$ be an open rod with an associated homotopy $(\bar{\boldsymbol{r}}, \bar{\boldsymbol{d}}_1)$, and let $\bar{\boldsymbol{d}}_2$ and $\bar{\boldsymbol{d}}_3$ be constructed from $\bar{\boldsymbol{r}}$ and $\bar{\boldsymbol{d}}_1$ according to (3.3). Let $\phi, \psi, \theta : [0, M] \times [0, 1] \to \mathbb{R}$ be the Euler-angle representation of $(\bar{\boldsymbol{d}}_1, \bar{\boldsymbol{d}}_2, \bar{\boldsymbol{d}}_3)$. Then*

$$R(\boldsymbol{r}, \boldsymbol{d}_1) = \int_0^L [\dot{\phi}(s, 1) + \dot{\psi}(s, 1)]\, ds.$$

From this lemma we conclude

COROLLARY 3.14. *For open rods $(\boldsymbol{r}, \boldsymbol{d}_1) \in \mathcal{A}^2_{r_0}$, $R(\boldsymbol{r}, \boldsymbol{d}_1)$ is independent of the choice of extension $(\bar{r}, \bar{\boldsymbol{d}}_1)$.*

PROOF OF LEMMA 3.13. Using the definitions of the Euler angles, we find

$$\partial_\lambda \bar{\boldsymbol{d}}_1(s, \lambda) \cdot \bar{\boldsymbol{d}}_2(s, \lambda) = \partial_\lambda \phi(s, \lambda) + \partial_\lambda \psi(s, \lambda)\, \cos\theta(s, \lambda).$$

For $s = 0, L$ and $\lambda \in [0, 1]$ we have set $\bar{\boldsymbol{d}}_3(s, \lambda) = \boldsymbol{e}_3$, and hence $\theta(s, \lambda) = 0$ for $s = 0, L$; therefore $\partial_\lambda \bar{\boldsymbol{d}}_1(s, \lambda) \cdot \bar{\boldsymbol{d}}_2(s, \lambda) = \partial_\lambda \phi(s, \lambda) + \partial_\lambda \psi(s, \lambda)$ for $s = 0, L$.

Since $\phi + \psi$ is a continuously differentiable function on $V := [0, L] \times [0, 1]$, the integral of the tangential derivative of $\phi + \psi$ along ∂V vanishes:

$$\oint_{\partial V} \frac{\partial}{\partial \tau}(\phi + \psi) = 0,$$

where τ is the clockwise-pointing unit vector tangential to ∂V.

Hence

$$\int_0^L [\partial_s\phi(s,1)+\partial_s\psi(s,1)]\,ds - \int_0^L [\partial_s\phi(s,0)+\partial_s\psi(s,0)]\,ds =$$
$$= \int_0^1 [\partial_\lambda\phi(L,\lambda)+\partial_\lambda\psi(L,\lambda)]\,d\lambda - \int_0^1 [\partial_\lambda\phi(0,\lambda)+\partial_\lambda\psi(0,\lambda)]\,d\lambda$$
$$= \int_0^1 [\partial_\lambda\phi(s,\lambda)+\partial_\lambda\psi(s,\lambda)]\,d\lambda\Big|_{s=0}^{s=L}$$
$$= \int_0^1 \partial_\lambda\bar{\boldsymbol{d}}_1(s,\lambda)\cdot\bar{\boldsymbol{d}}_2(s,\lambda)\,d\lambda\Big|_{s=0}^{s=L} = R(\boldsymbol{r},\boldsymbol{d}_1).$$

□

Within this framework Euler-angle formulae for open link, twist and writhe are obtained:

LEMMA 3.15. *Let* $(\boldsymbol{r},\boldsymbol{d}_1)$ *be an open rod in* $\mathcal{A}^2_{r_0}$. *Then*

$$Wr^o(\boldsymbol{r}) = \frac{1}{2\pi}\int_0^L \dot{\psi}(s)(1-\cos\theta(s))\,ds, \tag{3.16}$$

and

$$Tw^o(\boldsymbol{r},\boldsymbol{d}_1) = \frac{1}{2\pi}\int_0^L [\dot{\phi}(s)+\dot{\psi}(s)\cos\theta(s)]\,ds. \tag{3.17}$$

PROOF. The formula for twist is easily found by using (3.15) in the definition of twist, as in the proof of Lemma 3.13. For the writhe we apply Corollary 3.10 and use the fact that $v = \boldsymbol{e}_3$. □

The main result of this section states that for rods in $\mathcal{A}^2_{r_0}$ end rotation is equal to the open link:

THEOREM 3.16. *Let* $(\boldsymbol{r},\boldsymbol{d}_1)$ *be an open rod in* $\mathcal{A}^2_{r_0}$. *Then*

$$R(\boldsymbol{r},\boldsymbol{d}_1) = 2\pi Lk^o(\boldsymbol{r},\boldsymbol{d}_1) =$$
$$\int_0^L [\dot{\phi}(s)+\dot{\psi}(s)]\,ds = \phi(L)+\psi(L)-\phi(0)-\psi(0). \tag{3.18}$$

PROOF. By Corollary 3.4 and Lemma 3.15 we obtain

$$Lk^o(\boldsymbol{r},\boldsymbol{d}_1) = Wr^o(\boldsymbol{r}) + Tw^o(\boldsymbol{r},\boldsymbol{d}_1) = \frac{1}{2\pi}\int_0^L [\dot{\phi}(s)+\dot{\psi}(s)]\,ds.$$

Since $\int_0^L [\dot{\phi}(s)+\dot{\psi}(s)]\,ds = R$ by Lemma 3.13 we have the desired result. □

3.4. Critique of the approach

The example of Fig. 3.1 shows that end rotation can only be defined for a deformation history. For the purpose of analysis of elastic structures this dependence on deformation history is undesirable. The approach in this study, which is shared by many others (see the next section), is therefore to construct a class of deformation histories (homotopies) within which the end rotation *can* be expressed in terms of the initial and final states only. In this section we critically review the essential ingredients of this approach.

The closure. We obtain a separation into deformation classes by the introduction of a closure. The classes are characterised by the link of the closed rod-closure combination. The price we pay with this construction is the dependence on the choice of closure, which at first sight might seem to be a defect of the formulation. As Alexander & Antman [4] point out, however, this dependence is entirely natural: the precise form of the closure can be regarded as describing the way the rod is supported. Different systems of supports necessarily allow different classes of deformations.

Although in most applications throwing the rod around the clamp as illustrated in Fig. 3.1(d) is physically prevented, there do exist exceptions to this rule. In some recent experiments DNA molecules are manipulated with the help of a magnetic bead attached to the end of the molecule and held in a magnetic trap which allows the simultaneous application of a force and a moment [130]. In this case no rigid mechanical support is present and the molecule is free to loop around the beaded end. If this happens repeatedly, then the end can rotate over an arbitrarily large angle under the applied moment without a concomitant change in configuration (*cf*. Fig. 3.1: after a rotation of the end by -4π the rod returns to its original shape). Thus one might estimate the wrong energy, by $\pm 4\pi M$ for each 'looping', if a deformation would go outside the class, i.e., if link was not conserved. However, Rossetto & Maggs [116] in a recent paper show that for micron-sized beads the applied tension in many experiments is large enough (on the order of femtonewtons) to make these link violations rare, and they proceed to introduce a closure to study configurations of constant link.

Incidentally, the fact that a rotation of $\pm 4\pi$, and not one of $\pm 2\pi$, brings one back to where one started has its origin in the topological nature of SO(3), the group of rotations in $\mathbb{R}^3$. Specifically, SO(3) is not simply connected: for every rotation $R \in \mathrm{SO}(3)$ there are two homotopy classes of paths from the identity of the group to R. This means that for

a given rod orientation there are two distinct classes of configurations for the rod which cannot be deformed into each other while keeping the ends fixed. Rods with any even number of end turns (including zero) lie in one class; rods with any odd number of turns lie in the other. The same topological property forms the basis of the famous Dirac Belt Trick [74], which in classrooms is often illustrated by rotating a cup, held in the palm of one's hand, twice around a vertical axis by a suitable motion of the arm to bring both cup (with contents) and arm back to their initial positions (see [44] for a demonstration).

The reference curve. The class $\mathcal{A}^1_{\boldsymbol{r}_0}$ is defined for a fixed closed reference curve $\boldsymbol{r}_0$ (which may or may not include the unstressed centreline of the rod). The open writhe Wr^o will, in general, depend on the choice of this curve; on the one hand, by the fact that the reference curve restricts the class of admissible homotopies via the non-opposition condition, and on the other hand, by the explicit dependence on $\boldsymbol{t}_0$ in (3.12). Similarly, the end rotation R will depend on $\boldsymbol{r}_0$, as is to be expected since R is defined (in Definition 3.12) as the end rotation incurred in deforming $(\boldsymbol{r}_0, \bar{\boldsymbol{d}}_1(\cdot, 0))|_{[0,L]}$ into $(\boldsymbol{r}, \boldsymbol{d}_1)$.

For certain cases, however, the dependence can be described more precisely, as in Theorem 3.11 for straight reference curves. Though beyond the scope of this study, it is also possible to prove a similar result under rotation of more general reference curves. For instance, one could extend Definition 3.12, Lemma 3.13 and Corollary 3.14 to a larger class than $A^2_{\boldsymbol{r}_0}$ by requiring of $\boldsymbol{r}_0|_{[0,L]}$ only that its end tangents be equal. A complication would arise, however, in that violation of the non-opposition condition would no longer coincide with the Euler-angle singularity. Consequently, the non-opposition condition would no longer assure us of C^1-regularity of ϕ, ψ, and θ.

The non-opposition condition cannot be dropped. The non-opposition condition listed in the definition of $\mathcal{A}^1_{\boldsymbol{r}_0}$ (condition 5) is imposed by the application of Theorem 3.6. This might seem to be merely a technical restriction: after all, the open writhe is defined in terms of the writhe of the closed curve (Definition 3.3) and the latter is well-defined even if, somewhere along the homotopy, the non-opposition condition is violated. Therefore it might be expected that the statement of well-posedness holds true without condition 6 (even though our proof evidently does not), and that only non-self-intersection of the closed structure is required.

In fact the situation is not that simple. Figure 3.3 shows homotopy paths connecting the reference configuration (a) with the deformed rod-closure combinations (b), (c) and (d), where the rod itself (represented by the thick line) is the same in each of the three deformed states. We can imagine the deformed rod to be nearly planar, with the two strands crossing at a short distance from each other. Then the rod and its closure in case (b) have writhe close to -1.[1] In case (c) one adds or subtracts 1 to the writhe of the rod-closure combination for each full turn of the end. The writhe of the combination can therefore be made arbitrarily large. In case (d), finally, the writhe is close to -2.

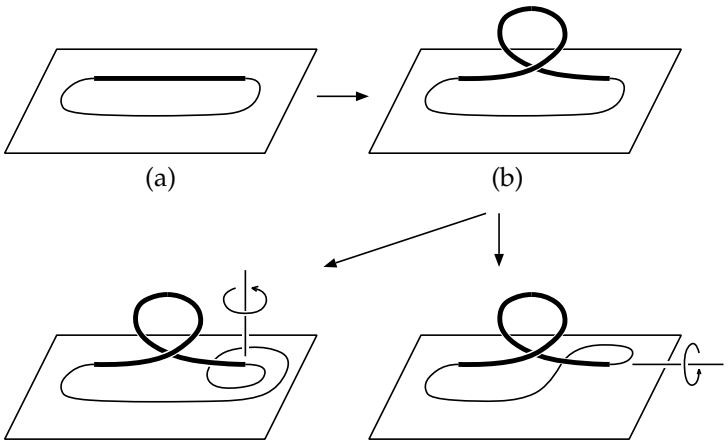


Figure 3.3: An example showing that the non-opposition condition 5 in Definition 3.2 cannot be disposed of. The three final states (b), (c) and (d), with identical shapes for the open rod, have different values of writhe. In going from (b) to either (c) or (d) the non-opposition condition is violated.

It is not difficult to see that one may construct a homotopy between (a) and (b) without violating the non-opposition condition, provided the loop has been twisted through an angle strictly less than π. Since the continuation homotopies to (c) and (d) satisfy all conditions of Definition 3.2 other than the non-opposition condition, it follows from Theorem 3.5 that these homotopies cannot be constructed without violating the non-opposition condition. This may also be verified by inspection.

[1] This may be verified by using the characterisation of writhe as the average of the directional writhing number, as explained in Section 3.2. This number is determined by counting signed crossings in a projection of the curve onto a plane.

This example shows that simply removing condition 5 from Definition 3.2 leads to ambiguities in the definition of writhe (and therefore of link). The example also suggests that if a well-defined writhe is to be constructed without the inclusion of the non-opposition condition, then additional restrictions must be imposed on the closure. In homotopy (c) the closure remains planar throughout the homotopy, but the end tangents vary; in homotopy (d) the end tangents are constant, but the closure is only planar at the beginning and the end of the homotopy. To rule out homotopies (c) and (d) (necessary for a well-defined writhe) we can require the end tangents to be fixed and the closure to be planar throughout the homotopy. It is possible that for a well-defined writhe further conditions must be imposed.

Euler-angle singularity *vs.* the non-opposition condition. When the reference configuration is straight and end tangents remain constant during the homotopy, the non-opposition condition is equivalent to avoidance of the Euler-angle singularity at $\theta = \pi$. Although this is partly a coincidence, the two issues both stem from the topological properties of S^2.

The Euler-angle singularity results from the fact that S^2 is not homeomorphic to (any part of) $\mathbb{R}^2$. Any parametrization of S^2 by a single cartesian coordinate system will therefore have at least one singular point. On the other hand, the non-opposition condition is necessary—in this article—for the single-integral representation of writhe of Theorem 3.6. In this representation the ambiguity of area 'enclosed' by a curve on S^2 is resolved by taking a perturbation approach. The non-opposition condition is the realization of the unavoidable limits of this approach, and therefore again stems from the topology of S^2.

As mentioned above, however, the non-opposition condition remains an unsatisfactory element in the definition of open writhe. Perhaps a concept of open writhe is possible that bypasses this condition.

Open writhe is not rotation invariant. The definition of $\mathcal{A}^1_{\boldsymbol{r}_0}$ depends on the choice of the reference configuration. For a given reference configuration, an open rod in $\mathcal{A}^1_{\boldsymbol{r}_0}$ may not be freely rotated without leaving $\mathcal{A}^1_{\boldsymbol{r}_0}$. This is readily demonstrated by rotating the reference configuration itself: after a rotation of π about an axis perpendicular to the plane of the reference curve the non-opposition condition is violated at every point on the curve.

This may lead to surprising results. In Figure 3.4 two homotopies are shown. The first is a variation on homotopy (b) of Figure 3.3, while in the second we lengthen the open-rod part and shorten the closure

part. In addition, we construct the homotopies such that the final configurations are close, up to a rotation (emphasized by the mark at one end of the open rod). In (a) the open writhe is close to 1, while in (b) it is close to 0.

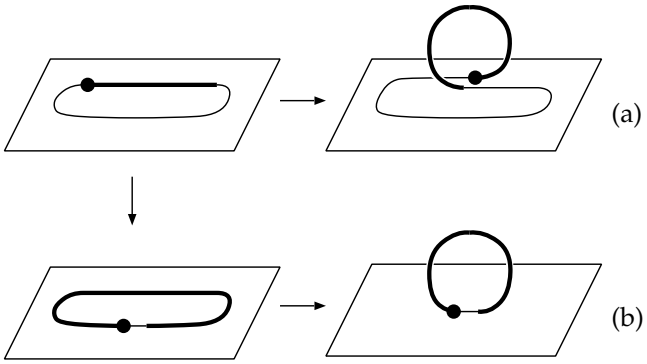


Figure 3.4: Two elements of $\mathcal{A}^1_{r_0}$ that differ only by a rotation, but for which the writhe is different. The dot emphasizes the difference in orientation.

This remark also resolves an issue raised in Remark 3.8: does the open writhe change continuously into the classical writhe for closed rods, when an open rod is transformed into a closed rod by lining up and connecting the ends? The answer is no—for the resulting closed writhe would be rotation-invariant, contradicting the remark above.

3.5. Discussion

The topological issues associated with large deformation discussed in this chapter are not of great concern in more traditional engineering applications. As long as deformations are such that the integral in (3.2) remains well-defined for a suitable choice of Euler angles (i.e., as long as the angles stay away from the polar singularity) open link and end rotation are given by (3.2). However, in more modern applications of structural mechanics, such as in molecular biology, large deformations occur more routinely and more care is required. Indeed, Fuller's 1971 paper [50] was inspired by supercoiling DNA molecules. In this study the author also already introduces a (planar) closing curve in order to compute the writhe of a simple (infinitely long) helix. Following the pioneering work of Fuller an extensive literature has emerged on the

application of elastic rod theory to DNA supercoiling (*cf*. the survey article by Schlick [122]).

Open rods have become popular models for DNA molecules since by the early 1990s single-molecule experiments have become possible. First this involved an applied force only [126]; later, once the molecule could be prevented from swivelling at its (magnetically) loaded end, this involved both an applied force and an applied moment [130]. Analytical studies have addressed DNA in isolation [10, 43] as well as in dilute solution using a statistical mechanics approach [95, 140, 99, 12].

Benham [10] appears to have been be the first one to write down an isoperimetric variational problem based on (3.2) in order to find equilibrium configurations subject to constant link. He first uses Fuller's result (3.9) with $\boldsymbol{r}_2$ taken to be a suitable closed planar curve. Then $Wr(\boldsymbol{r}_2) = 0$, and one obtains a single-integral formula for the writhe of the curve in question $\boldsymbol{r}_1$. When combined with the single integral for the twist (3.4), this leads, via (3.1), to a single-integral expression for the link, which in a suitable Euler-angle representation is given by (3.2). Benham then observes that "the integral expressions for *Lk*, *Tw* and *Wr* may be constructed regardless of whether the structure is closed". Many workers have since followed Benham's example to write link and writhe of an open rod as single integrals.

The implicit assumption in this approach is that there exists a continuous deformation from $\boldsymbol{r}_1$ to the planar curve $\boldsymbol{r}_2$ which avoids opposition of corresponding tangents. This may not be obvious, since the non-opposition condition must be applied to the closed combination of rod and suitable closure. In Section 3.4 we have given examples of what can go wrong if the condition is not satisfied. In addition, the example of Figure 3.4 shows that even within the limits of the non-opposition condition the writhe is not invariant under rotation, implying that this approach may lead to counterintuitive results.

Frequently, the end rotation needed in an energy analysis is simply assumed to be equal to the link as given by (3.2) (e.g., [43, 99]). We have shown that end rotation can only be defined in a consistent way within a class of homotopies of rods. Such a class is constructed with the help of a closure. We define end rotation independent of link and show the two to be equal within a suitable class of allowed deformations. We should also remark that our closure is a rod $(\boldsymbol{r}, \boldsymbol{d}_1)$ rather than just a centreline $\boldsymbol{r}$. Most authors initially only introduce a closed centreline, which makes the writhe well-defined, and subsequently assume the

closure to be twistless if a link or end rotation is required. We formalise this practise by explicitly specifying a $\boldsymbol{d}_1$ for the entire closed structure.

Knowing the precise restrictions on allowed deformations is important in statistical mechanics studies. To obtain the correct averages one must consider ensembles of admissible configurations. In numerical computations this means that one must take care to simulate configurations (through a Monte Carlo algorithm, a 'growth' algorithm or otherwise) with the right topological constraints. Specifically, one wants the configurations to be unknotted and to have constant link (although it is good to remember that DNA in its natural environment functions in the presence of topology-changing enzymes). This means that one must forbid self-crossings of the configuration as well as crossings of an (imaginary) closure. Mindful of this, the authors in [140] graft the ends of the molecule to an external surface and run 'sticks' out from the ends of the molecule to infinity, thereby 'virtually closing' the generated chain at infinity. A similar construction is used in [116]. In this latter work knotted configurations are not eliminated, it being argued that the molecular statistics is dominated by unknotted configurations.

In order to avoid the awkward non-opposition condition in Fuller's second theorem there have been direct approaches via the double integral (3.6) instead. In [129, 102] simple shapes are considered with planar closures for which the integral can be evaluated explicitly. It is then shown that the closure gives a relative contribution to the writhe which tends to zero as the length of the rod tends to infinity. The double integral is also used in the numerical study in [140], where it is shown that the contribution to Wr from the interaction of the closure with the basic chain is of the order of 1%. Various numerical schemes for the computation of the writhe double integral for a discretized curve are discussed and compared in [78]. Useful rigorous error bounds on numerically computed values of Wr based on polygonal (i.e., piece-wise linear) approximation are given in [23].

Our approach to a consistent definition of link, writhe and end rotation is firmly based on the introduction of a closure. There have been various formulations of writhe for open curves without the use of a closure. One approach, especially taken in knot theory and in studies of self-avoiding chains, is based on the characterisation of writhe as the average over all planar projections of the sum of signed crossings [50] (e.g., [106, 1]). This approach does not require a closure and can be applied to curves with arbitrary end tangents. It could form the basis of an alternative extension of (3.1) to open rods. (Here we can remark that

link also has an interpretation in terms of crossing numbers, namely: the linking number of two curves is equal to the number of all signed crossings of the curves in a regular planar projection, i.e., one satisfying a transversality condition; see [114].) The precise relation between the writhe of an open curve obtained via these planar projections and the open writhe obtained by using the tantrix area on S^2, or the double integral (3.6), is still an open problem.

If an exact writhe is not required and the fractional part modulo 2 is sufficient then formula (3.8) in terms of the area enclosed by the tantrix on the unit sphere can be used. Rossetto & Maggs [116] point out that this approach can also be used to generalise the writhe to curves whose end tangents are not aligned and therefore have open tantrices. By exploiting the connection between writhe and a geometric phase they show that the canonical way to close the curve is by means of a geodesic (great circle). This geodesic is unique as long as the two end points are not antipodal. The fractional writhe is then again given by the enclosed area. This prescription is used by Starostin [129] to derive results for the writhe of smooth as well as polygonal curves. Cantarella [23] generalises Fuller's second theorem, and with it the spherical area formula (3.8), to polygonal curves.

3.6. Proof of Theorem 3.11, and an extension

Let $f : S^2 \to \mathbb{R}$ be the function mentioned in the assertion. Pick $\boldsymbol{v}_0 \in \Omega$ and let Ω_0 be the connected component of Ω containing $\boldsymbol{v}_0$. Define the set

$$A = \{\boldsymbol{v} \in \Omega_0 : f(\boldsymbol{v}) = f(\boldsymbol{v}_0)\}.$$

The function f is continuous on Ω_0, implying that the set A is relatively closed in Ω_0. We will show below that f is constant on all open balls $B \subset \Omega_0$, implying that A is also open. Since A is non-empty it follows that $A = \Omega_0$ and the lemma is proved.

For a given vector $\omega \in S^2$, let R_ϕ denote the rotation around ω through an angle ϕ. We fix the direction of rotation in the following way: with respect to an orthonormal basis $(\omega, w, \omega \times w)$ for a suitable $w \in S^2$, write R_ϕ as

$$R_\phi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}.$$

With this choice,

$$\frac{d}{d\phi} R_\phi \boldsymbol{v}|_{\phi=0} = \boldsymbol{\omega} \times \boldsymbol{v} \qquad \text{for any } \boldsymbol{v} \in S^2. \tag{3.19}$$

Set $\boldsymbol{v}_\phi = R_\phi \boldsymbol{v}$. Using equation (3.19), we have

$$\frac{d}{d\phi} \frac{\boldsymbol{v}_\phi \times \boldsymbol{t}}{1 + \boldsymbol{v}_\phi \cdot \boldsymbol{t}} \cdot \dot{\boldsymbol{t}} = \frac{[(1 + \boldsymbol{v}_\phi \cdot \boldsymbol{t})((\boldsymbol{\omega} \times \boldsymbol{v}_\phi) \times \boldsymbol{t}) - (\boldsymbol{v}_\phi \times \boldsymbol{t})((\boldsymbol{\omega} \times \boldsymbol{v}_\phi) \cdot \boldsymbol{t})]}{(1 + \boldsymbol{v}_\phi \cdot \boldsymbol{t})^2} \cdot \dot{\boldsymbol{t}}. \tag{3.20}$$

Setting $\boldsymbol{\gamma} = |\boldsymbol{\omega} \times \boldsymbol{v}_\phi|$ we introduce an orthonormal coordinate system

$$e_1 = \boldsymbol{v}_\phi, \quad e_2 = \boldsymbol{\gamma}^{-1}\, \boldsymbol{\omega} \times \boldsymbol{v}_\phi, \quad e_3 = \boldsymbol{\gamma}^{-1}\, \boldsymbol{v}_\phi \times (\boldsymbol{\omega} \times \boldsymbol{v}_\phi),$$

and we write $\boldsymbol{t}_1$, $\boldsymbol{t}_2$, $\boldsymbol{t}_3$ for the coordinates of $\boldsymbol{t}$ with respect to this basis; these are functions of the curve parameter s. The right-hand side of (3.20) becomes

$$\boldsymbol{\gamma}\, \frac{(1 + \boldsymbol{t}_1)(\boldsymbol{t}_3 \dot{\boldsymbol{t}}_1 - \boldsymbol{t}_1 \dot{\boldsymbol{t}}_3) - \boldsymbol{t}_2(\boldsymbol{t}_2 \dot{\boldsymbol{t}}_3 - \boldsymbol{t}_3 \dot{\boldsymbol{t}}_2)}{(1 + \boldsymbol{t}_1)^2}.$$

Using the equalities $\boldsymbol{t}_1^2 + \boldsymbol{t}_2^2 + \boldsymbol{t}_3^2 = 1$ and $\boldsymbol{t}_1 \dot{\boldsymbol{t}}_1 + \boldsymbol{t}_2 \dot{\boldsymbol{t}}_2 + \boldsymbol{t}_3 \dot{\boldsymbol{t}}_3 = 0$ this is seen to be equal to

$$-\boldsymbol{\gamma}\, \frac{d}{ds}\, \frac{\boldsymbol{t}_3}{1 + \boldsymbol{t}_1}.$$

Therefore

$$\frac{d}{d\phi} \int \frac{\boldsymbol{v}_\phi \times \boldsymbol{t}}{1 + \boldsymbol{v}_\phi \cdot \boldsymbol{t}} \cdot \dot{\boldsymbol{t}}\, ds = -\boldsymbol{\gamma} \int \frac{d}{ds}\, \frac{\boldsymbol{t}_3}{1 + \boldsymbol{t}_1}\, ds = 0.$$

The last equality results from the assumption of aligned end tangents. This proves the theorem. □

It proves not to be possible to replace $\boldsymbol{t}_1(0)$ by an arbitrary constant vector $\boldsymbol{v} \in S^2$ in the above integral. Here we prove that within an appropriate neighbourhood of $\boldsymbol{t}_1(0)$ the integral (3.14) remains invariant under such a replacement.

We now have the following theorem.

THEOREM 3.17. *Let $\boldsymbol{a}, \boldsymbol{b}$ be two closed curves in $C^1_\#([0, L]; S^2)$, and R_ϕ a rotation over an angle ϕ as above. Assume $R_{\phi'} \boldsymbol{a}(s) \cdot R_{-\phi'} \boldsymbol{b}(s) > -1$ for all $0 \le \phi' \le \phi, s \in [0, L]$. Then*

$$\int \frac{R_\phi \boldsymbol{a} \times R_{-\phi} \boldsymbol{b}}{1 + R_\phi \boldsymbol{a} \cdot R_{-\phi} \boldsymbol{b}} \cdot (R_\phi \dot{\boldsymbol{a}} + R_{-\phi} \dot{\boldsymbol{b}}) = \int \frac{\boldsymbol{a} \times \boldsymbol{b}}{1 + \boldsymbol{a} \cdot \boldsymbol{b}} (\dot{\boldsymbol{a}} + \dot{\boldsymbol{b}}).$$

PROOF OF THEOREM 3.17. It is sufficient to look at integrals of the form

$$\frac{d}{d\phi}\int \frac{R_\phi \boldsymbol{a}\times \boldsymbol{b}}{1+R_\phi \boldsymbol{a}\cdot \boldsymbol{b}}\cdot \dot{\boldsymbol{b}}, \tag{3.21}$$

since

$$\begin{aligned}
&\frac{d}{d\phi}\int \frac{R_\phi \boldsymbol{a}\times R_{-\phi}\boldsymbol{b}}{1+R_\phi \boldsymbol{a}\cdot R_{-\phi}\boldsymbol{b}}\cdot (R_\phi \dot{\boldsymbol{a}}+R_{-\phi}\dot{\boldsymbol{b}})\\
&=\frac{d}{d\phi}\int \frac{R_\phi \boldsymbol{a}\times R_{-\phi}\boldsymbol{b}}{1+R_\phi \boldsymbol{a}\cdot R_{-\phi}\boldsymbol{b}}\cdot R_\phi \dot{\boldsymbol{a}}+\frac{d}{d\phi}\int \frac{R_\phi \boldsymbol{a}\times R_{-\phi}\boldsymbol{b}}{1+R_\phi \boldsymbol{a}\cdot R_{-\phi}\boldsymbol{b}}\cdot R_{-\phi}\dot{\boldsymbol{b}}\\
&=\frac{d}{d\phi}\int \frac{\boldsymbol{a}\times R_{-2\phi}\boldsymbol{b}}{1+\boldsymbol{a}\cdot R_{-2\phi}\boldsymbol{b}}\cdot \dot{\boldsymbol{a}}+\frac{d}{d\phi}\int \frac{R_{2\phi}\boldsymbol{a}\times \boldsymbol{b}}{1+R_{2\phi}\boldsymbol{a}\cdot \boldsymbol{b}}\cdot \dot{\boldsymbol{b}}\\
&=-\frac{d}{d\phi}\int \frac{R_{-2\phi}\boldsymbol{b}\times \boldsymbol{a}}{1+\boldsymbol{a}\cdot R_{-2\phi}\boldsymbol{b}}\cdot \dot{\boldsymbol{a}}+\frac{d}{d\phi}\int \frac{R_{2\phi}\boldsymbol{a}\times \boldsymbol{b}}{1+R_{2\phi}\boldsymbol{a}\cdot \boldsymbol{b}}\cdot \dot{\boldsymbol{b}}\\
&=\frac{d}{d\phi}\int \frac{R_{2\phi}\boldsymbol{b}\times \boldsymbol{a}}{1+\boldsymbol{a}\cdot R_{2\phi}\boldsymbol{b}}\cdot \dot{\boldsymbol{a}}+\frac{d}{d\phi}\int \frac{R_{2\phi}\boldsymbol{a}\times \boldsymbol{b}}{1+R_{2\phi}\boldsymbol{a}\cdot \boldsymbol{b}}\cdot \dot{\boldsymbol{b}}.
\end{aligned}\tag{3.22}$$

We will show that each of the arguments of these two integrals can be written as a derivative with respect to arclength plus an extra term. These two extra terms then sum up to a second derivative. Integrating over closed curves then means that the sum of these two integrals vanishes.

Recall that ω is the axis of rotation of R_ϕ. Using eq. (3.19) and differentiating (3.21) around $\phi = 0$, we get

$$\begin{aligned}
&\frac{d}{d\phi}\int \frac{R_\phi \boldsymbol{a}\times \boldsymbol{b}}{1+R_\phi \boldsymbol{a}\cdot \boldsymbol{b}}\cdot \dot{\boldsymbol{b}}|_{\phi=0}\\
&\quad=\int \frac{(1+\boldsymbol{a}\cdot \boldsymbol{b})((\omega\times \boldsymbol{a})\times \boldsymbol{b})\cdot \dot{\boldsymbol{b}}-(\boldsymbol{a}\times \boldsymbol{b})\cdot \dot{\boldsymbol{b}}((\omega\times \boldsymbol{a})\cdot \boldsymbol{b})}{(1+\boldsymbol{a}\cdot \boldsymbol{b})^2}\\
&\quad=\int \frac{(1+\boldsymbol{a}\cdot \boldsymbol{b})(\omega\times \boldsymbol{a})\cdot (\boldsymbol{b}\times \dot{\boldsymbol{b}})-\boldsymbol{a}\cdot (\boldsymbol{b}\times \dot{\boldsymbol{b}})((\omega\times \boldsymbol{a})\cdot \boldsymbol{b})}{(1+\boldsymbol{a}\cdot \boldsymbol{b})^2}.
\end{aligned}\tag{3.23}$$

Assume that ω is not parallel to $\boldsymbol{a}(s)$ for any $s\in[0,L]$. Then $(\boldsymbol{a}, \omega\times \boldsymbol{a}, \boldsymbol{a}\times(\omega\times \boldsymbol{a}))$ is a local orthogonal (but not necessarily orthonormal) basis. Denote these three vectors by $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3$ respectively. Write $\boldsymbol{b} =$

$b_1\gamma_1 + b_2\gamma_2 + b_3\gamma_3$ and $\dot{b} = \beta_1\gamma_1 + \beta_2\gamma_2 + \beta_3\gamma_3$. Then

$$b_1^2|\gamma_1|^2 + b_2^2|\gamma_2|^2 + b_3^2|\gamma_3|^2 = 1, \tag{3.24}$$

$$b_1\beta_1|\gamma_1|^2 + b_2\beta_2|\gamma_2|^2 + b_3\beta_3|\gamma_3|^2 = 0, \tag{3.25}$$

$$|\gamma_1| = 1, \text{ and } |\gamma_2| = |\gamma_3|, \tag{3.26}$$

$$b \times \dot{b} = \begin{pmatrix} \frac{|\gamma_2||\gamma_3|}{|\gamma_1|}(b_2\beta_3 - b_3\beta_2) \\ \frac{|\gamma_1||\gamma_3|}{|\gamma_2|}(b_3\beta_1 - b_1\beta_3) \\ \frac{|\gamma_1||\gamma_2|}{|\gamma_3|}(b_1\beta_2 - b_2\beta_1) \end{pmatrix}. \tag{3.27}$$

In this basis, (3.23) becomes

$$\int \frac{(1+\gamma_1\cdot b)\gamma_2\cdot(b\times\dot{b}) - \gamma_1\cdot(b\times\dot{b})(\gamma_2\cdot b)}{(1+\gamma_1\cdot b)^2}$$

which, using (3.27) equals

$$\begin{aligned}
&\int \frac{(1+b_1)|\gamma_2|^2\frac{|\gamma_1||\gamma_3|}{|\gamma_2|}(b_3\beta_1 - b_1\beta_3)}{(1+b_1)^2} \\
&\quad - \int \frac{|\gamma_2|^2 b_2|\gamma_1|\cdot\frac{|\gamma_2||\gamma_3|}{|\gamma_1|}(b_2\beta_3 - b_3\beta_2)}{(1+b_1)^2} \\
&\quad = |\gamma_2||\gamma_3|\int \frac{|\gamma_1|(1+b_1)(b_3\beta_1 - b_1\beta_3) - |\gamma_2|^2 b_2(b_2\beta_3 - b_3\beta_2)}{(1+b_1)^2}.
\end{aligned}$$

With (3.24), (3.25) and (3.26) this reduces further to

$$\begin{aligned}
&|\gamma_2||\gamma_3|\int \frac{b_3\beta_1 - b_1\beta_3 + \beta_3(-b_1^2 - |\gamma_2|^2 b_2^2) + b_3(b_1\beta_1 + |\gamma_2|^2 b_2\beta_2)}{(1+b_1)^2} \\
&\quad = |\gamma_2||\gamma_3|\int \frac{b_3(-|\gamma_3|^2 b_3\beta_3) + b_3\beta_1 - b_1\beta_3 + \beta_3(-1 + |\gamma_3|^2 b_3^2)}{(1+b_1)^2} \\
&\quad = |\gamma_2||\gamma_3|\int \frac{-\beta_3(1+b_1) + b_3\beta_1}{(1+b_1)^2}.
\end{aligned}$$

Using (3.26) once more we finally arrive at

$$\frac{d}{d\phi}\int \frac{R_\phi a\times b}{1+R_\phi a\cdot b}\cdot\dot{b}|_{\phi=0} = |\gamma_3|^2\int \frac{-\beta_3(1+b_1) + b_3\beta_1}{(1+b_1)^2}. \tag{3.28}$$

Now we work our way up from the other side. Let us consider

$$\frac{d}{dt}\Big(\frac{\gamma_3\cdot b}{1+a\cdot b}\Big), \tag{3.29}$$

where, as before, $\boldsymbol{\gamma}_3 = \boldsymbol{a} \times (\omega \times \boldsymbol{a})$. Performing the differentiation we obtain

$$\begin{aligned}
&\frac{(1+\boldsymbol{a}\cdot\boldsymbol{b})\frac{d}{dt}(\boldsymbol{\gamma}_3\cdot\boldsymbol{b}) - (\boldsymbol{\gamma}_3\cdot\boldsymbol{b})\frac{d}{dt}(\boldsymbol{a}\cdot\boldsymbol{b})}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2} \\
&= \frac{(1+\boldsymbol{a}\cdot\boldsymbol{b})(\boldsymbol{\gamma}_3\cdot\dot{\boldsymbol{b}}) - (\boldsymbol{\gamma}_3\cdot\boldsymbol{b})(\boldsymbol{a}\cdot\dot{\boldsymbol{b}})}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2} \\
&\qquad + \frac{(1+\boldsymbol{a}\cdot\boldsymbol{b})(\dot{\boldsymbol{\gamma}}_3\cdot\boldsymbol{b}) - (\boldsymbol{\gamma}_3\cdot\boldsymbol{b})(\dot{\boldsymbol{a}}\cdot\boldsymbol{b})}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2} \\
&= \frac{|\boldsymbol{\gamma}_3|^2((1+\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{\beta}_3 - \boldsymbol{b}_3\boldsymbol{\beta}_1)}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2} \\
&\qquad + \frac{(1+\boldsymbol{a}\cdot\boldsymbol{b})(\dot{\boldsymbol{\gamma}}_3\cdot\boldsymbol{b}) - (\boldsymbol{\gamma}_3\cdot\boldsymbol{b})(\dot{\boldsymbol{a}}\cdot\boldsymbol{b})}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}. \qquad (3.30)
\end{aligned}$$

Note that the first term is equal to the right-hand side of (3.28) up to a minus sign, since $\boldsymbol{a} = \boldsymbol{\gamma}_1$, and $\boldsymbol{\gamma}_1 \cdot \boldsymbol{b} = \boldsymbol{b}_1$. Using $\dot{\gamma}_3 = \boldsymbol{a}(\omega\cdot\dot{\boldsymbol{a}}) + \dot{\boldsymbol{a}}(\omega\cdot\boldsymbol{a})$, the numerator of the second term of (3.30) becomes

$$\begin{aligned}
&(1+\boldsymbol{a}\cdot\boldsymbol{b})(\dot{\boldsymbol{\gamma}}_3\cdot\boldsymbol{b}) - (\boldsymbol{\gamma}_3\cdot\boldsymbol{b})(\dot{\boldsymbol{a}}\cdot\boldsymbol{b}) \\
&\quad = -(1+\boldsymbol{a}\cdot\boldsymbol{b})(\boldsymbol{a}(\omega\cdot\dot{\boldsymbol{a}}) + \dot{\boldsymbol{a}}(\omega\cdot\boldsymbol{a}))\cdot\boldsymbol{b} - (\dot{\boldsymbol{a}}\cdot\boldsymbol{b})(\boldsymbol{a}\times(\omega\times\boldsymbol{a}))\cdot\boldsymbol{b} \\
&\quad = (\dot{\boldsymbol{a}}\cdot\boldsymbol{b})\big(-(\omega\cdot\boldsymbol{a})(1+\boldsymbol{a}\cdot\boldsymbol{b}) - \omega\cdot\boldsymbol{b} + (\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot\boldsymbol{a})\big) \\
&\qquad - (1+\boldsymbol{a}\cdot\boldsymbol{b})(\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot\dot{\boldsymbol{a}}) \\
&\quad = (\dot{\boldsymbol{a}}\cdot\boldsymbol{b})\big(-\omega\cdot\boldsymbol{a} - \omega\cdot\boldsymbol{b}\big) - (1+\boldsymbol{a}\cdot\boldsymbol{b})(\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot\dot{\boldsymbol{a}}) \\
&\quad = -(\dot{\boldsymbol{a}}\cdot\boldsymbol{b})\omega\cdot(\boldsymbol{a}+\boldsymbol{b}) - (\boldsymbol{a}\cdot\boldsymbol{b})(1+\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot\dot{\boldsymbol{a}}).
\end{aligned}$$

Reinstating the denominator $(1+\boldsymbol{a}\cdot\boldsymbol{b})^2$ we have

$$\begin{aligned}
\frac{d}{dt}\Big(\frac{\boldsymbol{\gamma}_3\cdot\boldsymbol{b}}{1+\boldsymbol{a}\cdot\boldsymbol{b}}\Big) = &\\
&\frac{|\boldsymbol{\gamma}_3|^2((1+\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{\beta}_3 - \boldsymbol{b}_3\boldsymbol{\beta}_1)}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2} \\
&\quad + \frac{-(\dot{\boldsymbol{a}}\cdot\boldsymbol{b})\omega\cdot(a+\boldsymbol{b}) - (a\cdot\boldsymbol{b})(1+\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot\dot{\boldsymbol{a}})}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}.
\end{aligned}$$

We can perform the same calculations assuming that $(\boldsymbol{b}, \omega\times\boldsymbol{b}, \boldsymbol{b}\times(\omega\times\boldsymbol{b}))$ is a basis. Write this basis as $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\delta}_3)$, and set $a = \boldsymbol{a}_1\boldsymbol{\delta}_1 + \boldsymbol{a}_2\boldsymbol{\delta}_2 + \boldsymbol{a}_3\boldsymbol{\delta}_3$ and $\dot{\boldsymbol{a}} = \boldsymbol{\alpha}_1\boldsymbol{\delta}_1 + \boldsymbol{\alpha}_2\boldsymbol{\delta}_2 + \boldsymbol{\alpha}_3\boldsymbol{\delta}_3$. When we add to (3.29) its

symmetric equivalent, we get

$$\begin{aligned}
&\frac{d}{dt}\Big(\frac{\boldsymbol{\gamma}_3\cdot\boldsymbol{a}}{1+\boldsymbol{a}\cdot\boldsymbol{b}}\Big)+\frac{d}{dt}\Big(\frac{\boldsymbol{\delta}_3\cdot\boldsymbol{b}}{1+\boldsymbol{a}\cdot\boldsymbol{b}}\Big)\\
&\quad=\frac{|\boldsymbol{\gamma}_3|^2((1+\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{\beta}_3-\boldsymbol{b}_3\boldsymbol{\beta}_1)}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}\\
&\qquad-\frac{(\dot{\boldsymbol{a}}\cdot\boldsymbol{b})\omega\cdot(\boldsymbol{a}+\boldsymbol{b})-(\boldsymbol{a}\cdot\boldsymbol{b})(1+\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot\dot{\boldsymbol{a}})}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}\\
&\qquad+\frac{|\boldsymbol{\delta}_3|^2((1+\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{\alpha}_3-\boldsymbol{a}_3\boldsymbol{\alpha}_1)}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}-\\
&\qquad-\frac{(\dot{\boldsymbol{b}}\cdot\boldsymbol{a})\omega\cdot(\boldsymbol{a}+\boldsymbol{b})-(\boldsymbol{a}\cdot\boldsymbol{b})(1+\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot\dot{\boldsymbol{b}})}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}\\
&\quad=\frac{|\boldsymbol{\gamma}_3|^2((1+\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{\beta}_3-\boldsymbol{b}_3\boldsymbol{\beta}_1)}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}\\
&\qquad+\frac{|\boldsymbol{\delta}_3|^2((1+\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{\alpha}_3-\boldsymbol{a}_3\boldsymbol{\alpha}_1)}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}\\
&\qquad-\frac{[\frac{d}{dt}(\boldsymbol{a}\cdot\boldsymbol{b})]\omega\cdot(\boldsymbol{a}+\boldsymbol{b})-(\boldsymbol{a}\cdot\boldsymbol{b})(1+\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot(\dot{\boldsymbol{a}}+\dot{\boldsymbol{b}}))}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}.
\end{aligned}\tag{3.31}$$

For the last term in (3.31) we finally derive

$$\begin{aligned}
&\frac{-[\frac{d}{dt}(\boldsymbol{a}\cdot\boldsymbol{b})]\omega\cdot(\boldsymbol{a}+\boldsymbol{b})-(\boldsymbol{a}\cdot\boldsymbol{b})(1+\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot(\dot{\boldsymbol{a}}+\dot{\boldsymbol{b}}))}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}\\
&=\frac{(1+\boldsymbol{a}\cdot\boldsymbol{b})(\omega\cdot(\dot{\boldsymbol{a}}+\dot{\boldsymbol{b}}))-[\frac{d}{dt}(1+\boldsymbol{a}\cdot\boldsymbol{b})]\omega\cdot(\boldsymbol{a}+\boldsymbol{b})}{(1+\boldsymbol{a}\cdot\boldsymbol{b})^2}-\omega\cdot(\dot{\boldsymbol{a}}+\dot{\boldsymbol{b}})\\
&=\frac{d}{dt}\Big(\frac{\omega\cdot(\boldsymbol{a}+\boldsymbol{b})}{1+\boldsymbol{a}\cdot\boldsymbol{b}}\Big)-\frac{d}{dt}(\omega\cdot(\boldsymbol{a}+\boldsymbol{b}))\\
&=\frac{d}{dt}\Big(\frac{-(\boldsymbol{a}\cdot\boldsymbol{b})\omega\cdot(\boldsymbol{a}+\boldsymbol{b})}{1+\boldsymbol{a}\cdot\boldsymbol{b}}\Big).
\end{aligned}\tag{3.32}$$

Now we can wrap things up: the two integrals in (3.22) are symmetric equivalents, and using derivative one in (3.29) and derivative two in

(3.32), we have

$$\begin{aligned}
&\frac{d}{d\phi}\int \frac{R_\phi \boldsymbol{b}\times \boldsymbol{a}}{1+\boldsymbol{a}\cdot R_\phi \boldsymbol{b}}\cdot \dot{\boldsymbol{a}}|_{\phi=0} + \frac{d}{d\phi}\int \frac{R_\phi \boldsymbol{a}\times \boldsymbol{b}}{1+R_\phi \boldsymbol{a}\cdot \boldsymbol{b}}\cdot \dot{\boldsymbol{b}}|_{\phi=0}\\
&\quad = \int \frac{d}{dt}\Big(\frac{\boldsymbol{\gamma}_3\cdot \boldsymbol{a}}{1+\boldsymbol{a}\cdot \boldsymbol{b}}\Big) + \int \frac{d}{dt}\Big(\frac{\boldsymbol{\delta}_3\cdot \boldsymbol{b}}{1+\boldsymbol{a}\cdot \boldsymbol{b}}\Big) + \int \frac{d}{dt}\Big(\frac{-(\boldsymbol{a}\cdot \boldsymbol{b})\omega\cdot(\boldsymbol{a}+\boldsymbol{b})}{1+\boldsymbol{a}\cdot \boldsymbol{b}}\Big)\\
&\quad = 0.
\end{aligned}$$

We have thus far assumed that $(\boldsymbol{a}, \omega, \boldsymbol{a}\times(\omega\times\boldsymbol{a}))$ and $(\boldsymbol{b}, \omega\times\boldsymbol{b}, \boldsymbol{b}\times(\omega\times\boldsymbol{b}))$ are both bases. We can easily generalize to all $\omega \in S^2$, i.e., to all ω for which both $(\boldsymbol{a}, \omega, \boldsymbol{a}\times(\omega\times\boldsymbol{a}))$ and $(\boldsymbol{b}, \omega\times\boldsymbol{b}, \boldsymbol{b}\times(\omega\times\boldsymbol{b}))$ are not bases. Note that in this case ω has to be parallel to either $\boldsymbol{a}(s)$ or $\boldsymbol{b}(s)$ for some $s\in[0,L]$ (on S^2 this means $\omega = \pm\boldsymbol{a}(s)$ or $\omega = \pm\boldsymbol{b}(s)$). For given C^1 curves $\boldsymbol{a}(s)$ and $\boldsymbol{b}(s)$, the set of vectors parallel to $\boldsymbol{a}(s)$ and $\boldsymbol{b}(s)$ for some $s\in[0,L]$ has zero measure on S^2. Since the last integral in (3.23) is a continuous functional in ω which vanishes on a set of ω's of measure 4π, it has to vanish for all $\omega\in S^2$. This shows that for small enough ϕ, by Taylor's theorem, there exists a constant C such that

$$\begin{aligned}
&\int \frac{R_\phi \boldsymbol{a}\times R_{-\phi}\boldsymbol{b}}{1+R_\phi \boldsymbol{a}\cdot R_{-\phi}\boldsymbol{b}}\cdot(R_\phi\dot{\boldsymbol{a}} + R_{-\phi}\dot{\boldsymbol{b}})\\
&= -\int \frac{R_{-2\phi}\boldsymbol{b}\times \boldsymbol{a}}{1+\boldsymbol{a}\cdot R_{-2\phi}\boldsymbol{b}}\cdot\dot{\boldsymbol{a}} + \int \frac{R_{2\phi}\boldsymbol{a}\times \boldsymbol{b}}{1+R_{2\phi}\boldsymbol{a}\cdot \boldsymbol{b}}\cdot\dot{\boldsymbol{b}} \qquad (3.33)\\
&\le \int \frac{\boldsymbol{a}\times \boldsymbol{b}}{1+\boldsymbol{a}\cdot \boldsymbol{b}}\cdot(\dot{\boldsymbol{a}}+\dot{\boldsymbol{b}}) + C\phi^2,
\end{aligned}$$

provided that the two integrals in eq. (3.33b) are well-defined for all $0\le\phi'\le\phi$, i.e., that

$$\begin{aligned}
\boldsymbol{a}\cdot R_{-2\phi'}\boldsymbol{b} > -1 \text{ for all } 0\le\phi'\le\phi,\\
R_{2\phi'}\boldsymbol{a}\cdot \boldsymbol{b} > -1 \text{ for all } 0\le\phi'\le\phi.
\end{aligned}$$

This is equivalent to

$$R_{\phi'}\boldsymbol{a}\cdot R_{-\phi'}\boldsymbol{b} > -1 \text{ for all } 0\le\phi'\le\phi.$$

For larger ϕ, set

$$f(\phi,\boldsymbol{a},\boldsymbol{b}) = \int \frac{R_\phi \boldsymbol{a}\times R_{-\phi}\boldsymbol{b}}{1+R_\phi \boldsymbol{a}\cdot R_{-\phi}\boldsymbol{b}}\cdot(R_\phi\dot{\boldsymbol{a}} + R_{-\phi}\dot{\boldsymbol{b}}).$$

Take two sufficiently small subangles ϕ_1 and ϕ_2. Then

$$\begin{aligned} f(\phi_1+\phi_2,\boldsymbol{a},\boldsymbol{b}) &= f(\phi_2, R_{\phi_1}\boldsymbol{a}, R_{-\phi_1}\boldsymbol{b}) \\ &= f(\phi_2, R_{\phi_1}\boldsymbol{a}, R_{-\phi_1}\boldsymbol{b}) - f(\phi_1,\boldsymbol{a},\boldsymbol{b}) + f(\phi_1,\boldsymbol{a},\boldsymbol{b}) \\ &= f(\phi_2, R_{\phi_1}\boldsymbol{a}, R_{-\phi_1}\boldsymbol{b}) \\ &\qquad - f(0, R_{\phi_1}\boldsymbol{a}, R_{-\phi_1}\boldsymbol{b}) + f(\phi_1,\boldsymbol{a},\boldsymbol{b}) \\ &\le C(\phi_2^2+\phi_1^2). \end{aligned}$$

Hence, if we partition ϕ into n small equal subangles, then

$$\left|\int \frac{R_\phi\boldsymbol{a}\times R_{-\phi}\boldsymbol{b}}{1+R_\phi\boldsymbol{a}\cdot R_{-\phi}\boldsymbol{b}}\cdot(R_\phi\dot{\boldsymbol{a}}+R_{-\phi}\dot{\boldsymbol{b}}) - \int\frac{\boldsymbol{a}\times\boldsymbol{b}}{1+\boldsymbol{a}\cdot\boldsymbol{b}}(\dot{\boldsymbol{a}}+\dot{\boldsymbol{b}})\right| \le Cn\frac{|\phi|^2}{n^2} \longrightarrow 0 \text{ as } n\longrightarrow\infty,$$

provided

$$R_{\phi'}\boldsymbol{a}\cdot R_{-\phi'}\boldsymbol{b} > -1 \text{ for all } 0\le\phi'\le\phi.$$

□

CHAPTER 4

Regularity of minimizers for a simple continuum model for lipid membranes

We study the minimization problem

Problem (A)

$$\inf\{H(u) \mid u \in K_m\},$$

where

$$H(u) := \int_{\mathbb{R}} \left[h(u) - \alpha u\, \kappa * u\right],$$

$$h(u) := (u + c_0)\log(u/c_0 + 1) - u$$

$$K_m := \left\{ u \in L^1(\mathbb{R}) \;\middle|\; \int_{\mathbb{R}} u = m,\ u \geq 0,\ u + \tau_h u \leq 1 \right\},$$

$$\tau_h u(x) := u(x-h), \quad \kappa(x) := \frac{e^{-|x|}}{2},$$

$$\alpha > 0, \quad c_0 \in \left(0, \frac{1}{2}\right), \quad h > 0.$$

The two sets where the inequality constraints are saturated, the *contact sets*, are defined as

$$\Omega_d := \{x \in \mathbb{R} \mid u(x) = 0 \ a.e.\}, \tag{4.1}$$

$$\Omega_u := \{x \in \mathbb{R} \mid u(x) + (\tau_h u)(x) = 1 \ a.e.\}. \tag{4.2}$$

The main aim of this chapter is to study the regularity of minimizers for Problem (A). A solid understanding of the regularity of minimizers is essential to later investigate the existence of minimizers for Problem (A), which is the subject of Chapter 5. In that same chapter, we also

study a related constrained minimization problem in which the upper contact condition $u + \tau_h u \leq 1$ is dropped:

Problem (B)

$$\inf \left\{ H(u) \mid u \in K_m^B \right\}, \tag{4.3}$$

where

$$K_m^B := \left\{ u \in L^1(\mathbb{R}) \;\middle|\; \int_{\mathbb{R}} u = m,\ u \geq 0 \right\},$$

and $H(u)$ as before.

Throughout this chapter, all integrals are over $\mathbb{R}$, unless stated otherwise. The complement of a set V is denoted by V^c.

4.1. Introduction

The set of admissible functions $K_m \subset L^1(\mathbb{R})$ consists of functions of little regularity, and boundedness of the functional H does not improve this; as a result, we have little *a priori* knowledge of the regularity of a minimizer. In variational problems of this type, but without constraints, some regularity can be proved for stationary points by using the Euler-Lagrange equation,

$$\log(1 + u/c_0) = 2\alpha\, \kappa * u.$$

A simple bootstrap argument—if $u \in W^{k,p}$, then $\kappa * u \in W^{k+2,p}$, and therefore $u \in W^{k+2,p}$—now provides u with C^∞-regularity.

For the problem at hand—minimization of H under this specific set of pointwise and non-local constraints—this argument fails, since the Euler-Lagrange equation is augmented with three Lagrange multipliers that intervene with this regularity bootstrap (see (5.5) in Chapter 5). We therefore use a different strategy, based on an estimate of the gap in Jensen's inequality in terms of the regularity of the underlying function.

Let $f : \mathbb{R} \to \mathbb{R}$ be convex, and choose $\rho \in L^1(\mathbb{R})$ with $\rho \geq 0$ and $\int \rho = 1$. Construct a regularization kernel by setting $\rho_\varepsilon(x) = 1/\varepsilon\, \rho(x/\varepsilon)$. A consequence of Jensen's inequality is that for any $u \in L^1(\mathbb{R})$

$$\int f(u) = \int \rho_\varepsilon * f(u) \geq \int f(\rho_\varepsilon * u).$$

With a formal argument we can relate the difference between the two sides in this inequality to the regularity of u. Writing $\rho_\varepsilon * u \approx u + c\varepsilon^2 u''$,

with $c = (1/2)\int y^2\rho(y)\,dy$, we find

$$\int [f(u) - f(\rho_\varepsilon * u)] \approx -c\varepsilon^2 \int f'(u)u'' = c\varepsilon^2 \int f''(u)u'^2,$$

suggesting that if $u \in H^1(\mathbb{R})$, then the difference on the left-hand side should be of order ε^2.

The convergence rate of ε^2 is connected to the regularity of the function u, as a simple example shows. Take $f(u) = u^2$ and consider a function with a jump singularity such as $u(x) = H(x) - H(x-1)$, where $H(x)$ is the Heaviside function, so that $u \notin H^1(\mathbb{R})$. Now choose as regularization kernels the functions $\rho_\varepsilon(x) = \frac{1}{2\varepsilon}(H(x+\varepsilon) - H(x-\varepsilon))$, after which an explicit calculation shows that the convergence is only of order ε:

$$\int [u^2 - u_\varepsilon^2] = \frac{2}{3}\varepsilon.$$

Our contribution to this topic consists of two parts. First, we prove both the suggestion made above and its converse. In both cases, $\rho_\varepsilon(x) = 1/\varepsilon\,\rho(x/\varepsilon)$ for some symmetric $\rho \in L^1(\mathbb{R})$ satisfying $\rho \geq 0$, $\int \rho = 1$, and $\int y^2\rho(y)\,dy < \infty$.

THEOREM 4.1. *Let $u \in H^1(\mathbb{R})$ and $f \in W^{2,\infty}(\mathbb{R})$. Then*

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon^2} \int_{\mathbb{R}} [f(u) - f(\rho_\varepsilon * u)] = c \int_{\mathbb{R}} f''(u)u'^2,$$

where

$$c = \frac{1}{2}\int_{\mathbb{R}} y^2\rho(y)\,dy.$$

THEOREM 4.2. *Let $f : \mathbb{R} \to \mathbb{R}$ satisfy $f'' \geq c_1 > 0$ in the sense of distributions. Then there exists a constant C such that any $u \in L^1(\mathbb{R})$ for which the limit*

$$M := \limsup_{\varepsilon\to 0} \frac{1}{\varepsilon^2} \int_{\mathbb{R}} [f(u) - f(\rho_\varepsilon * u)] \tag{4.4}$$

is finite satisfies $\|u\|_{H^1}^2 \leq \frac{CM}{c_1}$.

The proofs of both of these theorems consist of a reduction to the well-known characterization of Sobolev norms by Bourgain, Brezis, and Mironescu [13]. They are given in Lemmas 4.7 and 4.8.

Our second contribution consists in applying Theorem 4.2 to the constrained minimizers of H to obtain an *a priori* regularity result. We thus show an alternative way of establishing regularity of minimizers,

which is especially useful when the Euler-Lagrange equation is unavailable or does not provide this information.

A key element is that K_m is closed under convolution with ρ_ε: if $u \in K_m$, then $u_\varepsilon := \rho_\varepsilon * u \in K_m$, so that if u is a minimizer, then $H(u) \leq H(u_\varepsilon)$ for all $\varepsilon > 0$.

The function h satisfies the differential inequality of this theorem, with $c_1 = 1/(1 + c_0)$. From the estimate

$$\left| \int [u_\varepsilon \, \kappa * u_\varepsilon - u \, \kappa * u] \right| \leq C\varepsilon^2 \int u^2 \leq C\varepsilon^2 \left\| u \right\|_{L^1(\mathbb{R})},$$

it then follows that

$$0 \leq H(u_\varepsilon) - H(u) \leq \int [h(u_\varepsilon) - h(u)] + O(\varepsilon^2).$$

By Theorem 4.2 we conclude that $u \in H^1(\mathbb{R})$. The details of this argument are given in Section 4.2.

4.2. Minimizers have H^1-regularity

In this section we prove Theorems 4.1 and 4.2, and show that minimizers of Problems (A) and (B) are elements of $H^1(\mathbb{R})$.

THEOREM 4.3. *Let u be a minimizer of Problem (A) or of Problem (B). Then $u \in H^1(\mathbb{R})$. In particular, u is continuous.*

PROOF. As explained in the introduction, the regularity of a constrained minimizer u follows from comparing the energy levels of u and $u_\varepsilon := \rho_\varepsilon * u$ and using Theorem 4.1. The argument differs only slightly between Problems (A) and (B).

The second term of H can be estimated as follows,

LEMMA 4.4. *There exists a constant $C \in \mathbb{R}$ such that for all $u \in L^2(\mathbb{R})$*

$$0 \leq \int_{\mathbb{R}} u \, \kappa * u - \int_{\mathbb{R}} u_\varepsilon \, \kappa * u_\varepsilon \leq C\varepsilon^2 \int_{\mathbb{R}} u^2.$$

PROOF OF LEMMA 4.4. Let $\hat{u}$ be the Fourier transform of u, defined by $\hat{u}(k) := \int_{\mathbb{R}} e^{-ixk} u(x) dx$. Note that $\hat{\kappa}(k) = 1/(1 + k^2)$. Therefore,

$$\int_{\mathbb{R}} (\rho_\varepsilon * u) \, \kappa * (\rho_\varepsilon * u) = \int_{\mathbb{R}} |\hat{u}(k)|^2 |\hat{\rho}(\varepsilon k)|^2 \frac{1}{1 + k^2} \, dk. \tag{4.5}$$

Since ρ is symmetric, positive, and has finite second moment, $\hat{\rho}$ is twice differentiable in 0 and satisfies $|\hat{\rho}| \leq \int \rho = 1$, $\hat{\rho}(0) = 1$, and $\hat{\rho}'(0) =$

0. Therefore there exists an $\alpha > 0$ such that $|\hat{\rho}(k)|^2 \geq 1 - \alpha k^2$ for k sufficiently small. Using (4.5), we now have

$$\begin{aligned} 0 &\leq \int_{\mathbb{R}} u\, \kappa * u - \int_{\mathbb{R}} (\rho_\varepsilon * u)\, \kappa * (\rho_\varepsilon * u) x \qquad (4.6) \\ &= \int_{\mathbb{R}} |\hat{u}(k)|^2 \frac{1}{1+k^2} \left(1 - |\hat{\rho}(\varepsilon k)|^2\right) dk \\ &\leq \int_{|k| \leq \frac{1}{\varepsilon}} |\hat{u}(k)|^2 \frac{\alpha \varepsilon^2 k^2}{1+k^2}\, dk + \int_{|k| > \frac{1}{\varepsilon}} |\hat{u}(k)|^2 \frac{1}{1+k^2}\, dk \\ &\leq \alpha \varepsilon^2 \int_{\mathbb{R}} |\hat{u}(k)|^2\, dk + \frac{\varepsilon^2}{1+\varepsilon^2} \int_{\mathbb{R}} |\hat{u}(k)|^2\, dk \\ &= \varepsilon^2 (\alpha + 1) \int_{\mathbb{R}} u^2(x)\, dx. \qquad (4.7) \end{aligned}$$

□

Since both K_m and K_m^B are closed under convolution with ρ_ε, u_ε is admissible, and therefore

$$0 \leq H(u_\varepsilon) - H(u) \leq \int [h(u_\varepsilon) - h(u)] + C\varepsilon^2 \int u^2.$$

For Problem (A) the last term can be estimated by

$$C\varepsilon^2 \|u\|_{L^1(\mathbb{R})} \|u\|_{L^\infty(\mathbb{R})} \leq Cm\varepsilon^2,$$

where the second inequality follows from combining both inequality constraints. Therefore $\int [h(u) - h(u_\varepsilon)] = O(\varepsilon^2)$, from which the result follows from an application of Theorem 4.2.

For Problem (B) the proof is concluded similarly with the L^∞-bound given by the next lemma. □

LEMMA 4.5. *Any minimizer u of Problem (B) satisfies $u \in L^\infty(\mathbb{R})$.*

PROOF. Define the set $A = \{x \in \mathbb{R} : u(x) > \frac{1}{2} \|u\|_{L^\infty(\mathbb{R})}\}$. Perturbation with functions v such that $\operatorname{supp} v \subset A$ and $\int v = 0$ gives the equation

$$\log(u/c_0 + 1) - 2\alpha\, \kappa * u = \lambda \quad \text{a.e. on } A,$$

where $\lambda \in \mathbb{R}$ is a Lagrange multiplier. Hence

$$1 + u(x)/c_0 = e^{\lambda + 2\alpha\, \kappa * u(x)} \quad \text{for a.e. } x \in A.$$

Since κ is bounded, $\kappa * u \in L^\infty(\mathbb{R})$, with

$$\|\kappa * u\|_{L^\infty(\mathbb{R})} \leq \|\kappa\|_{L^\infty(\mathbb{R})} \|u\|_{L^1(\mathbb{R})} = \frac{m}{2},$$

and therefore

$$\|u\|_{L^\infty(\mathbb{R})} = \|u\|_{L^\infty(A)} \leq c_0\big(e^{\lambda+\alpha m} - 1\big).$$

□

REMARK 4.6. With the continuity given by Theorem 4.3 the result of Lemma 4.5 can be supplemented with an explicit bound. Choosing a sequence $(x_n) \subset \mathbb{R}$ such that $u(x_n) > 0$ and $u(x_n) \to 0$, we can write

$$\lambda = \lim_{n\to\infty} \big[\log\big(1 + u(x_n)/c_0\big) - 2\alpha(\kappa * u)(x_n)\big] \leq 0,$$

so that

$$\|u\|_{L^\infty(\mathbb{R})} \leq c_0\big(e^{\alpha m} - 1\big). \tag{4.8}$$

PROOF OF THEOREM 4.2. As before, set $u_\varepsilon = \rho_\varepsilon * u$. First we prove that if (4.4) holds, then

$$\limsup_{\varepsilon\to 0} \frac{1}{\varepsilon^2} \int_{\mathbb{R}} (u^2 - u_\varepsilon^2) \leq \frac{2}{c_1} M. \tag{4.9}$$

The reduces the proof to the case $f(u) = u^2$. Expanding around $u_\varepsilon(x)$, we find by assumption that

$$\begin{aligned} f(u(x-y)) \geq f(u_\varepsilon(x)) &+ f'(u_\varepsilon(x))\,(u(x-y) - u_\varepsilon(x)) \\ &+ \frac{c_1}{2}\,(u(x-y) - u_\varepsilon(x))^2\,. \end{aligned} \tag{4.10}$$

Using the identity

$$\begin{aligned} \int_{\mathbb{R}} f'(u_\varepsilon(x))\,(u(x-y) - u_\varepsilon(x))\,\rho_\varepsilon(y)dy &= f'(u_\varepsilon(x))\,(u_\varepsilon(x) - u_\varepsilon(x)) \\ &= 0 \end{aligned}$$

we obtain from (4.10),

$$\begin{aligned} \int_{\mathbb{R}} f(u_\varepsilon(x))dx &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(u_\varepsilon(x))\rho_\varepsilon(y)dydx \\ &\leq \int_{\mathbb{R}}\int_{\mathbb{R}} f(u(x-y))\rho_\varepsilon(y)dydx \\ &\quad - \frac{c_1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}} (u(x-y) - u_\varepsilon(x))^2\,\rho_\varepsilon(y)dydx \\ &= \int_{\mathbb{R}} \rho_\varepsilon * f(u) - \frac{c_1}{2}\int_{\mathbb{R}} [\,\rho_\varepsilon * u^2 - u_\varepsilon^2] \\ &= \int_{\mathbb{R}} f(u) - \frac{c_1}{2}\int_{\mathbb{R}} [u^2 - u_\varepsilon^2]. \end{aligned}$$

This shows that

$$\begin{aligned}\limsup_{\varepsilon\to 0}\frac{1}{\varepsilon^2}\int_{\mathbb{R}}[u^2-u_\varepsilon^2] &\le \limsup_{\varepsilon\to 0}\frac{2}{c_1\varepsilon^2}\int_{\mathbb{R}}[f(u)-f(u_\varepsilon)]\\ &= \frac{2}{c_1}M,\end{aligned}$$

which proves (4.9). Now to prove the lemma for $f(u) = u^2$, define $\gamma_\varepsilon := \rho_\varepsilon * \rho_\varepsilon$. Then

$$\begin{aligned}\int_{\mathbb{R}} u_\varepsilon^2 &= \int_{\mathbb{R}} u\gamma_\varepsilon * u\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} u(x)u(y)\gamma_\varepsilon(x-y)dxdy\\ &= \int_{\mathbb{R}} u^2 - \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}}\left(u(x)-u(y)\right)^2\gamma_\varepsilon(x-y)dxdy.\end{aligned}$$

Hence

$$\limsup_{\varepsilon\to 0}\frac{1}{\varepsilon^2}\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{|u(x)-u(y)|^2}{|x-y|^2}|x-y|^2\gamma_\varepsilon^2(x-y)dxdy \le \frac{4}{c_1}M < \infty.$$

The sequence of functions defined by

$$\phi_\varepsilon(x) = \frac{x^2}{\varepsilon^2}\gamma_\varepsilon(x)$$

is a regularization kernel, and hence by Theorem 2 in [13] we conclude $u \in H^1(\mathbb{R})$ and that there exists a constant K such that

$$\|u\|_{H^1}^2 \le \frac{4M}{c_1K}.$$

This completes the proof. □

The proof of Theorem 4.1 is split into the following two lemmas. Recall that $u_\varepsilon = \rho_\varepsilon * u$. Then

$$\int_{\mathbb{R}}[f(u)-f(u_\varepsilon)] = -\int_{\mathbb{R}} f'(u_\varepsilon)(u_\varepsilon-u) + R_\varepsilon,$$

where

$$|R_\varepsilon| \le \left\|f''\right\|_{L^\infty}\|u_\varepsilon-u\|_{L^2}^2.$$

LEMMA 4.7. *Let $u \in H^1(\mathbb{R})$, and ρ_ε a regularization kernel as before. Then*

$$\|u_\varepsilon-u\|_{L^2}^2 = o(\varepsilon^2).$$

PROOF.

$$\begin{aligned}\int_{\mathbb{R}}(u_\varepsilon - u)^2 dx &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}}\left(u(x-y)-u(x)\right)\rho_\varepsilon(y)dy\right)^2 dx \\ &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}}\left[-\int_0^1 u'(x-sy)yds\right]\rho_\varepsilon(y)dy\right)^2 dx \\ &= \int_{\mathbb{R}}\left(\int_0^1\int_{\mathbb{R}} u'(x-z)\frac{z}{s}\rho_\varepsilon\left(\frac{z}{s}\right)dzds\right)^2 dx. \qquad (4.11)\end{aligned}$$

Set $\sigma(y) = y\rho(y)$, and

$$\sigma_\varepsilon(z,s) = \frac{1}{\varepsilon s}\sigma\left(\frac{z}{\varepsilon s}\right) = \frac{z}{\varepsilon^2 s^2}\rho\left(\frac{z}{\varepsilon s}\right) = \frac{z}{\varepsilon s^2}\rho_\varepsilon\left(\frac{z}{s}\right).$$

Then (4.11) becomes

$$\begin{aligned}&\varepsilon^2\int_{\mathbb{R}}\left(\int_0^1\int_{\mathbb{R}} u'(x-z)\sigma_\varepsilon(z,s)dzds\right)^2 dx \\ &\qquad\qquad \le \varepsilon^2\int_0^1\int_{\mathbb{R}}\left(\int_{\mathbb{R}} u'(x-z)\sigma_\varepsilon(z,s)dz\right)^2 dxds. \qquad (4.12)\end{aligned}$$

Now set

$$v_\varepsilon(x,s) = \left(\int_{\mathbb{R}} u'(x-z)\sigma_\varepsilon(z,s)dz\right).$$

We claim that $v_\varepsilon(x,s) \to 0$ as $\varepsilon \to 0$ for all $s \in (0,1]$ and for almost all $x \in \mathbb{R}$, and we claim that v_ε is uniformly bounded in L^2. By the dominated convergence theorem we then conclude that $v_\varepsilon \to 0$ in the L^2-norm, for all $s \in (0,1]$, which proves the lemma.

Regarding the first claim, note that, since ρ is symmetric, $\int_{\mathbb{R}} x\rho(x) = 0$. Using this, and the properties of $\sigma_\varepsilon(z,s)$, we get in the limit as $\varepsilon \to 0$, that

$$\begin{aligned}\int_{\mathbb{R}} u'(x-z)\sigma_\varepsilon(z,s)dz &\to u'(x)\frac{1}{s}\int_{\mathbb{R}}\sigma(y)dy \\ &= u'(x)\frac{1}{s}\int_{\mathbb{R}} y\rho(y)dy \\ &= 0.\end{aligned}$$

For the second claim, note that $\int_{\mathbb{R}} \sigma_\varepsilon(z,s)dz$ is constant for all $\varepsilon > 0$ and $s \in (0,1]$. Denote this integral by c. Then

$$\begin{aligned} v_\varepsilon^2(x,s) &\leq c\left(\int_{\mathbb{R}} u'^2(x-z)|\sigma_\varepsilon(z,s)|dz\right) \\ &\leq c\left\|(|\sigma_\varepsilon(\cdot,s)| * u'^2(\cdot))(x)\right\|_{L^1} \\ &\leq c\left\|\sigma_\varepsilon(\cdot,s)\right\|_{L^1}\left\|u'^2\right\|_{L^1}^2. \end{aligned}$$

This proves the claim. □

LEMMA 4.8. *Same assumptions as in Lemma 4.7. Assume furthermore that $f \in W^{2,\infty}(\mathbb{R})$. Then*

$$\lim_{\varepsilon\to 0} -\frac{1}{\varepsilon^2}\int_{\mathbb{R}} f'(u)(u_\varepsilon - u) \to c\int_{\mathbb{R}} f''(u)u'^2,$$

where

$$c = \frac{1}{2}\int_{\mathbb{R}} y^2\rho(y)dy.$$

PROOF.

$$\begin{aligned} -\int_{\mathbb{R}} f'(u)(u_\varepsilon - u) &= -\int_{\mathbb{R}}\int_{\mathbb{R}} f'(u(x))\left(u(x-y)-u(x)\right)\rho_\varepsilon(y)dydx \\ &= \int_0^1\int_{\mathbb{R}}\int_{\mathbb{R}} u'(z)yf'(u(z+sy))\rho_\varepsilon(y)dydzds \\ &= \int_0^1\int_{\mathbb{R}} u'(z)\left(\int_{\mathbb{R}} f'(u(z+sy))y\rho_\varepsilon(y)dy\right)dzds. \end{aligned}$$

Define for fixed s, $\sigma'(y) = y\rho(y)$, and $\sigma_\varepsilon(y) = \frac{1}{\varepsilon}\sigma(\frac{y}{\varepsilon})$. Then $\sigma_\varepsilon'(y) = \frac{1}{\varepsilon^2}\sigma'(\frac{y}{\varepsilon}) = \frac{y}{\varepsilon^3}\rho(\frac{y}{\varepsilon}) = \frac{1}{\varepsilon^2}y\rho_\varepsilon(y)$. Now, for any $g \in H^1$,

$$\begin{aligned} \frac{1}{\varepsilon^2}\int_{\mathbb{R}} y\rho_\varepsilon(y)g(x+y)dy &= \int_{\mathbb{R}} \sigma_\varepsilon'(y)g(x+y)dy \\ &= -\int_{\mathbb{R}} \sigma_\varepsilon(y)g'(x+y)dy \\ &= -\int_{\mathbb{R}} \sigma_\varepsilon(-y)g'(x-y)dy \\ &\to -g'(x)\int_{\mathbb{R}} \sigma \text{ in } L^2(\mathbb{R}). \end{aligned}$$

Writing $f'(u(z+sy)) = f'\left(u(s(\zeta+y))\right) =: g(\zeta+y)$, we proceed

$$\begin{aligned}
&\frac{1}{\varepsilon^2}\int_0^1\int_{\mathbb{R}} u'(z)\int_{\mathbb{R}} f'(u(z+sy))y\rho_\varepsilon(y)dydzds\\
&\quad=\int_0^1\int_{\mathbb{R}} u'(z)\int f'(u(z+sy))\sigma_\varepsilon'(y)dydzds\\
&\quad=\int_0^1 s\int_{\mathbb{R}} u'(s\zeta)\int_{\mathbb{R}} f'(u(s(\zeta+y))\sigma_\varepsilon'(y)dyd\zeta ds\\
&\quad=-\int_0^1 s^2\int_{\mathbb{R}} u'(s\zeta)\int_{\mathbb{R}} f''(u(s(\zeta+y)))u'(s(\zeta+y))\sigma_\varepsilon(y)dyd\zeta ds\\
&\quad=-\int_0^1 s^2\int_{\mathbb{R}} u'(s\zeta)[f''(u(s\cdot))u'(s\cdot)*\sigma_\varepsilon]d\zeta ds\\
&\quad\to\left(\int_{\mathbb{R}}\sigma(y)dy\right)\int_0^1 s^2\int_{\mathbb{R}} u'(s\zeta)f''(u(s\zeta))u'(s\zeta)d\zeta ds\quad\text{ as }\varepsilon\to 0\\
&\quad=\left(\int_{\mathbb{R}}\sigma(y)dy\right)\int_0^1 sds\int_{\mathbb{R}} u'(z)f''(u(z))u'(z)dz\\
&\quad=-\frac{1}{2}\left(\int_{\mathbb{R}}\sigma(y)dy\right)\int_{\mathbb{R}} u'(y)^2f''(u(y))dy.
\end{aligned}$$

We now complete the proof by showing that

$$\left|\int_{\mathbb{R}} u'(z)\int_{\mathbb{R}}\sigma_\varepsilon'(y)f'(u(z+sy))dydz\right|$$

is uniformly bounded for all $s\in(0,1]$. Set $\lambda_\varepsilon(s,\eta)=\sigma_\varepsilon(\eta/s)$. Then for all $s\in(0,1]$,

$$\begin{aligned}
&\left|\int_{\mathbb{R}} u'(z)\int_{\mathbb{R}}\sigma_\varepsilon'(y)f'(u(z+sy)dydz\right|\\
&\qquad\le\left\|u'\right\|_{L^2}\left\|\int_{\mathbb{R}} f'(u(z+sy))\sigma_\varepsilon'(y)dy\right\|_{L^2}\\
&\qquad=s\left\|u'\right\|_{L^2}\left\|\int_{\mathbb{R}} f''(u(z+sy))u'(z+sy)\sigma_\varepsilon(y)dy\right\|_{L^2}\\
&\qquad=s\left\|u'\right\|_{L^2}\left\|\int_{\mathbb{R}} f''(u(z+\eta))u'(z+\eta)\lambda_\varepsilon(s,\eta)d\eta\right\|_{L^2}\\
&\qquad\le\left\|u'\right\|_{L^2}^2\left\|f''\right\|_{L^\infty}.
\end{aligned}$$

We now conclude by the dominated convergence theorem that

$$-\frac{1}{\varepsilon^2}\int_{\mathbb{R}} f'(u)(u_\varepsilon-u)\to c\int_{\mathbb{R}} f''(u)u'^2,$$

with

$$c = -\frac{1}{2}\int_{\mathbb{R}} \sigma(y)dy = \frac{1}{2}\int_{\mathbb{R}} y^2\rho(y)dy.$$

□

CHAPTER 5

Existence of minimizers for a simple continuum model for lipid membranes

5.1. Introduction

We again study the minimization problem

Problem (A)

$$\inf\{H(u) \mid u \in K_m\},$$

where

$$\begin{aligned}
H(u) &:= \int_{\mathbb{R}} \big[h(u) - \alpha u\, \kappa * u\big], \\
h(u) &:= (u + c_0)\log(u/c_0 + 1) - u \\
K_m &:= \left\{ u \in L^1(\mathbb{R}) \;\middle|\; \int_{\mathbb{R}} u = m,\ u \geq 0,\ u + \tau_h u \leq 1 \right\}, \\
\tau_h u(x) &:= u(x-h), \quad \kappa(x) := \frac{e^{-|x|}}{2}, \\
\alpha &> 0, \quad c_0 \in \left(0, \frac{1}{2}\right), \quad h > 0.
\end{aligned}$$

The two sets where the inequality constraints are saturated, the *contact sets*, are defined as

$$\Omega_d := \{x \in \mathbb{R} \mid u(x) = 0 \ a.e.\}, \tag{5.1}$$

$$\Omega_u := \{x \in \mathbb{R} \mid u(x) + (\tau_h u)(x) = 1 \ a.e.\}. \tag{5.2}$$

In the previous chapter we have established in Theorem 4.3 that any minimizer of Problem (A) has H^1-regularity. This *a priori* information forms a key ingredient in the study of the existence of minimizers, the main topic of this chapter. In the process of this investigation, we also provide some insight into the structure of minimizing solutions. The results on both the existence of minimizers and their qualitative properties are complemented by numerical computations presented in Section 5.8.

As in the previous chapter, all integrals are over $\mathbb{R}$, unless stated otherwise, and the complement of a set V is denoted by V^c.

Convolution and non-local constraints

There are a number of ingredients in this minimization problem that make a proof of existence non-trivial. Among these is the unbounded domain, which is a possible source for lack of compactness of a minimizing sequence. However, the Concentration-Compactness Lemma by Lions [86], provides a potent tool to negotiate this possible lack of compactness. Two other ingredients provide greater obstacles: the convolution integral $\int u\,\kappa{*}u$, and the non-local inequality constraint $u + \tau_h u + 1 \leq 1$.

The convolution integral makes it difficult to control how the energy changes when we modify an admissible function. Especially if we add two functions—particularly useful in the Concentration-Compactness principle—it is often hard to estimate how the energy of the sum relates to the sum of the energies of the two functions. It proves crucial to show that minimizers have compact support to get around this problem: given two minimizers $u_1 \in K_{m_1}$ and $u_2 \in K_{m_2}$ with compact support, we may translate u_1 or u_2 such that $u_1 + u_2 \in K_{m_1+m_2}$. Estimation of the energy of the resulting function is now an easy matter.

The upper contact condition $u + \tau_h u \leq 1$ is perhaps the most notable feature of Problem (A), and is a direct consequence of our effort to model lipid bilayers consisting of lipid molecules with polar heads and apolar tails. Evidently, such a non-local condition makes it difficult to perform cut and paste arguments, which is one of the hallmarks of global minimization methods. To show how this contact condition alters the qualitative properties of minimizers, in Section 5.4 we briefly digress to study a simplified problem, in which this constraint has been dropped:

Problem (B)

$$\inf\left\{H(u) \mid u \in K_m^B\right\}, \tag{5.3}$$

where

$$K_m^B := \left\{u \in L^1(\mathbb{R}) \;\middle|\; \int_{\mathbb{R}} u = m,\ u \geq 0\right\},$$

and $H(u)$ as before.

There are two important differences between Problems (A) and (B): first, the Euler-Lagrange equations for minimizers of Problem (B) are simpler: they have one Lagrange multiplier less; second, for Problem (B) we can use symmetrization techniques, whilst for Problem (A) we cannot. A lemma by Riesz, discussed in detail in Section 5.4, states that the integral $\int u\,\kappa * u$ does not decrease if we replace u by its symmetrized equivalent u^* (defined in that same section). These two properties indicate that minimizers of Problem (B) have a simpler structure.

Existence of minimizers

The Concentration-Compactness Lemma gives three possibilities for any sequence in $L^1(\mathbb{R})$ with constant mass $\int u$: a subsequence either vanishes, splits, or is compact. If there exists a minimizing sequence for Problem (A) that is compact (in the sense of Concentration-Compactness), then there exists a minimizer [86]. Set

$$I(m) := \inf\{H(u) \mid u \in K_m\}.$$

Then, analogous to Theorem II.1 in [86], we prove in Theorem 5.13 that any minimizing sequence for Problem (A) with mass m is compact (in the sense of Concentration-Compactness) if and only if

$$I(m) < I(\mu) + I(m-\mu) \text{ for all } \mu \in (0, m). \tag{5.4}$$

Hence, given parameter values α, h, and c_0, existence of a minimizer is assured if (5.4) holds. The opposite is not necessarily true: if (5.4) does not hold for some $m > 0$, then there may still exist a minimizer for that value of m.

In proving existence for either minimization problems, we take the direct route by proving (5.4), rather than excluding dichotomy and vanishing for a minimizing sequence. We hence need quite some information on the properties of $I(m)$ as a function of m. In Section 5.3 we prove the following basic characteristics of $I(m)$:

- $I(m) \geq -\frac{\alpha m^2}{2}$.

- The energy of a vanishing minimizing sequence is zero. Hence $I(m) \leq 0$ for all $m > 0$ (Lemma 5.14).
- $I(m)$ is non-increasing in m (Lemma 5.16).

Furthermore, a key ingredient to prove (5.4) is to know that $I(m) < 0$ for all $m > 0$. In Lemma 5.26 we prove that if $2\alpha c_0 > 1$, for all $m > 0$ there exist admissible functions with negative energy, and therefore $I(m) < 0$ for all $m > 0$. If $2\alpha c_0 > 1$, however, we show in that same lemma that for small $m > 0$, there does not exist any admissible function with negative energy, and hence minimizing sequences exist that are not compact. This suggests that minimizers may not exist at all for small $m > 0$.

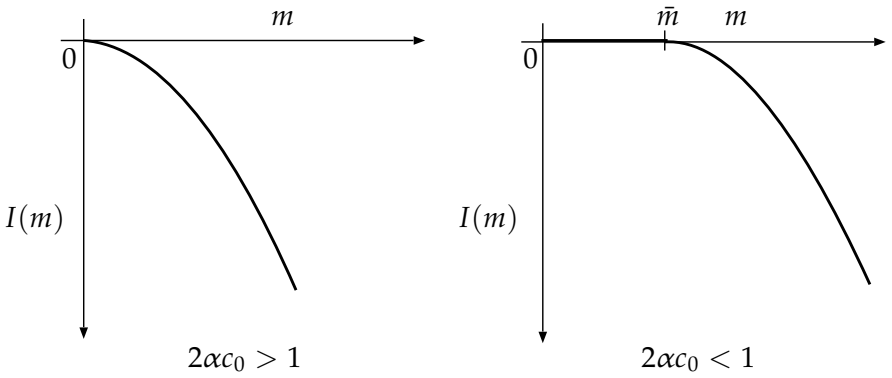


Figure 5.1: Two different cases of existence. If $2\alpha c_0 > 1$ (left), $I(m) < 0$ for all $m > 0$. This indicates (together with (5.4)) that minimizers exist for all $m > 0$. If $2\alpha c_0 < 1$ (right), $I(m) = 0$ for $m \in [0, \bar{m}]$ ($\bar{m} > 0$, and may depend on the other parameters), and $I(m) < 0$ for all $m > \bar{m}$. This at least means that not all minimizing sequences are compact, and may imply that minimizers do not exist at all for $m \in (0, \bar{m}]$. For Problem (B) we show that minimizers do exist for $m > \bar{m}$; for Problem (A) the above sketch is corroborated by numerical simulations, indicating that minimizers should also exist for all $m > \bar{m}$.

The picture that follows from these results is sketched in Figure 5.1. For both Problem (A) and (B) we prove that minimizers exist for all $m > 0$ provided that $2\alpha c_0 > 1$ (Theorems 5.28 and 5.32). For $2\alpha c_0 < 1$, the main method of proof to show that minimizers exist for sufficiently large m is by again proving (5.4). A sufficient condition for (5.4) to hold is to show that $I(x)/x$ is a strictly decreasing function at $x = m$. For Problem (B) this is relatively easy, resulting in Theorem 5.33.

For Problem (A), however, the upper contact condition does not allow one to perform the same argument. Instead, we have performed simulations to investigate whether minimizers exist for large m if $2\alpha c_0 < 1$, and how $\bar{m}$—the infimum of all m for which there exists a minimizer—depends on α and c_0. The numerical calculations are performed using a gradient flow, and are presented in Section 5.8. They suggest that also for Problem (A) minimizers exist for sufficiently large mass, and that the minimal amount of mass for which a minimizer exists increases with decreasing α.

The simulations also provide insight into the structure of minimizing solutions. In Figure 5.2 two numerical minimizers are plotted, for different parameter values. The main features they display are:

- the solutions are continuous (Theorem 4.3 in Chapter 4),
- they have compact support (Lemma 5.29 for $2\alpha c_0 > 1$),
- they are periodic on Ω_u with period $2h$ (see Lemma 5.37 for the precise statement),
- their upper contact sets Ω_u are connected,
- for large α or large h the upper and lower contact sets have a non-empty intersection, while for small α or small h they do not.

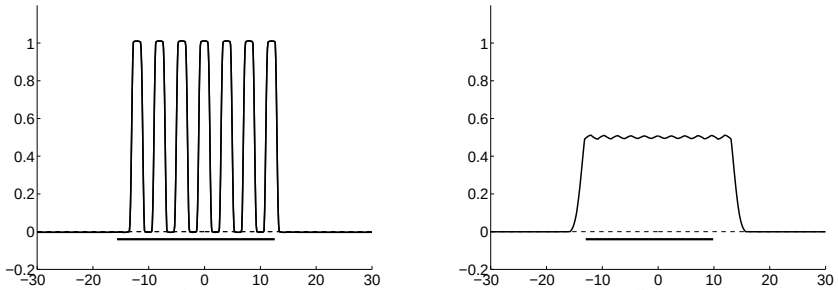


Figure 5.2: Two example minimizers of Problem (A), computed with the gradient flow algorithm described in Section 5.8. The horizontal line below the solution signifies the part of the solution where $u + \tau_h u = 1$. Note that in (a) there is a non-empty intersection of the lower and upper contact sets, Ω_d and Ω_u, while in (b) this intersection is empty.

The last two properties are only suggested by numerics, and remain unproven. Especially the case when Ω_u and Ω_d have a non-empty intersection forms a particularly interesting situation: on this intersection the two Lagrange multipliers μ and ν may both be non-zero, but have to balance each other. The simulations suggest that when α or h increases this intersection becomes larger. Below we state a conjecture

speculating on the existence and structure of a non-empty intersection of the contact sets. Given h and c_0, let $\bar{\alpha} > 0$ be the infimum over all α for which minimizers exists for at least one $m > 0$.

CONJECTURE 5.1. *Fix $h > 0$, $c_0 \in (0, 1/2)$ and $\bar{\alpha}$ as above. For $\alpha > \bar{\alpha}$, there exists $0 \leq \bar{m} < m_1 < \infty$ such that for all $m > \bar{m}$ Problem (A) has a minimizer with contact sets Ω_u^m and Ω_d^m. Moreover,*

(1) *u_m is symmetric, has compact support, and Ω_u^m is a (possibly empty) interval;*
(2)

$$\begin{aligned} \Omega_u^m \cap \Omega_d^m &= \emptyset \quad \text{for all } m \in (\bar{m}, m_1], \\ \Omega_u^m \cap \Omega_d^m &\neq \emptyset \quad \text{for all } m > m_1. \end{aligned}$$

We hope to return to this at a later stage.

5.2. The Euler-Lagrange equations

Having established that minimizers of Problems (A) and (B) are continuous in Chapter 4, the second major ingredient of a proof of existence of minimizers of Problem (A) is the derivation of the Euler-Lagrange equations. We first prove that the upper contact set Ω_u is compact.

LEMMA 5.2. *Let u be a minimizer of Problem (A) with upper contact set Ω_u. Then Ω_u is compact.*

PROOF. Evidently, Ω_u is compact if and only if $\Omega_u + \{0, h\}$ is compact. We prove by contradiction. Since Ω_u is closed, assume that Ω_u is unbounded. Then the set $\{u(x) > 1/2\} \cap (\Omega_u + \{0, h\})$ is also unbounded. Hence there exists a set $\{x_n\}$ such that $u(x_n) \geq 1/2$, and $x_n \to \infty$ as $n \to \infty$. By Theorem 4.3, $u \in H^1$. We now use that for any positive function $u \in H^1(\mathbb{R})$ there exists a constant $C \in \mathbb{R}$ such that

$$\|u\|_{L^\infty} \leq C \|u\|_{L^1}^{1/3} \|u'\|_{L^2}^{2/3}.$$

Then

$$\begin{aligned}\|u\|_{L^1(x_n,\infty)} &\geq C\frac{\|u\|^3_{L^\infty(x_n,\infty)}}{\|u'\|^2_{L^2(x_n,\infty)}}\\ &\geq C\frac{\|u\|^3_{L^\infty(x_n,\infty)}}{\|u'\|^2_{L^2(\mathbb{R})}}\\ &\geq \frac{C}{8\,\|u'\|^2_{L^2(\mathbb{R})}}\\ &> 0 \text{ for all } n \in \mathbb{N}.\end{aligned}$$

But $u \in L^1(\mathbb{R})$, and hence

$$\|u\|_{L^1(x,\infty)} \to 0 \text{ as } x \to \infty,$$

a contradiction. □

THEOREM 5.3. *Let u be a minimizer of Problem (A). Then there exist $\mu, \nu \in L^\infty(\mathbb{R})$ and $\lambda \in \mathbb{R}$ such that*

$$\log(u/c_0 + 1) - 2\alpha\,\kappa * u = \lambda + \mu - \nu - \tau_{-h}\nu. \tag{5.5}$$

In addition, μ and ν are positive and

$$\begin{aligned}\operatorname{supp}\mu &\subset \Omega_d,\\ \operatorname{supp}\nu &\subset \Omega_u.\end{aligned}$$

DEFINITION 5.4. *A function $u \in K_m$ is called a stationary point of Problem (A) if there exist $\mu, \nu \in L^\infty(\mathbb{R})$ and $\lambda \in \mathbb{R}$ such that*

$$\log(u/c_0 + 1) - 2\alpha\,\kappa * u = \lambda + \mu - \nu - \tau_{-h}\nu,$$

in the sense of distributions, with in addition, μ and ν positive, and

$$\operatorname{supp}\mu \subset \Omega_d, \quad \operatorname{supp}\nu \subset \Omega_u.$$

REMARK 5.5. Using exactly the same techniques as in the proof of the Euler-Lagrange equations for Problem (A) we can also derive the Euler-Lagrange equations for Problem (B). Let u be a minimizer for Problem (B) with contact set Ω_d. Then there exist $\lambda \in \mathbb{R}$ and $\mu \in L^\infty(\mathbb{R})$ such that

$$\log(u/c_0 + 1) - 2\alpha\,\kappa * u = \lambda + \mu.$$

Moreover, μ is positive, and $\operatorname{supp}\mu \subset \Omega_d$.
A function $u \in K^B_m$ is called a stationary point of Problem (B) if there exist $\lambda \in \mathbb{R}$ and $\mu \in L^\infty(\mathbb{R})$ such that

$$\log(u/c_0 + 1) - 2\alpha\,\kappa * u = \lambda + \mu,$$

and such that μ is positive and $\operatorname{supp}\mu \subset \Omega_d$.

The proof of Theorem 5.3 follows along the lines of Theorem 2.5 in Chapter 2. Let $X = L^1(\mathbb{R})$ and $Y = \mathbb{R}\times\mathbb{R}\times L^1(\mathbb{R})\times L^1(\mathbb{R})$. X' is identified with $L^\infty(\mathbb{R})$, and Y' with $\mathbb{R}\times\mathbb{R}\times L^\infty(\mathbb{R})\times L^\infty(\mathbb{R})$. We fix the function u, with contact sets Ω_d and Ω_u defined in (5.1), and introduce the cone of admissible perturbations V,

$$V := \{v \in X \,:\, \exists\{\varepsilon_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+, \\ \varepsilon_n \to 0 \text{ such that } B(u+\varepsilon_n v) \geq 0 \;\forall n \in \mathbb{N}\},$$

where

$$\begin{aligned} B: \quad X &\to Y, \\ u &\mapsto \left(\int u - m,\; m - \int u,\; u,\; 1-u-\tau_h u\right). \end{aligned}$$

LEMMA 5.6. *Let u be a minimizer. Then $H'(u)\cdot v \geq 0$ for all $v \in \overline{V}$.*

PROOF. For any $v \in V$, since u is a minimizer,

$$H(u+\varepsilon_n v) - H(u) \geq 0 \qquad \text{for all } n \in \mathbb{N}.$$

H is Fréchet differentiable in u so that

$$0 \leq H(u+\varepsilon_n v) - H(u) = \varepsilon_n H'(u)\cdot v + o(\varepsilon_n \|v\|_X),$$

from which it follows that $H'(u)\cdot v \geq 0$. Now, given any $v \in \overline{V}$, take a sequence $v_m \subset V$ that converges to v in X. Since $H'(u) : X \to \mathbb{R}$ is a continuous linear operator, $H'(u)\cdot v_m \to H'(u)\cdot v$. Hence $H'(u)\cdot v \geq 0$ for any $v \in \overline{V}$. □

Let

$$\begin{aligned} b: \quad X &\to Y, \\ v &\mapsto \left(\int v,\; -\int v,\; v,\; -v-\tau_h v\right) \end{aligned}$$

be the derivative of B. We then define the bounded linear operator

$$\begin{aligned} \beta: \quad X &\to Y, \\ v &\mapsto \left(\int v,\; -\int v,\; v\chi_{\Omega_d},\; (-v-\tau_h v)\chi_{\Omega_u}\right), \end{aligned}$$

where χ_M is the characteristic function of the set M. $\overline{V}$ can now be characterized in a more convenient way.

LEMMA 5.7. *For any $u \in K_m$,*

$$\overline{V} = W := \{v \in X \,:\, \beta v \geq 0\},$$

We postpone the proof to the end of this section.

$\overline{V}$ is a closed convex cone, with dual cone

$$\overline{V}^\perp = \{\gamma \in X' : \ \langle \gamma, v\rangle \geq 0 \quad \forall v \in \overline{V}\}.$$

For any $y \in Y$, denote its i-th component by y_i. Let

$$P = \{y \in Y : \ y_3 \geq 0 \text{ on } \Omega_d \text{ and } y_4 \geq 0 \text{ on } \Omega_u\}.$$

This also is a closed convex cone, with dual cone

$$P^\perp = \{f \in Y' : \ \langle f, y\rangle \geq 0 \quad \forall y \in P\}.$$

LEMMA 5.8. *If $(\lambda_1, \lambda_2, \mu, \nu) \in P^\perp$, then* $\operatorname{supp}\mu \subset \Omega_d$, $\operatorname{supp}\nu \subset \Omega_u$, *and $\mu, \nu \geq 0$.*

PROOF. Given any $y \in Y$ such that $\operatorname{supp} y_3 \subset \Omega_d^c$ and $\operatorname{supp} y_4 \subset \Omega_u^c$, $y \in P$ and $-y \in P$. Hence, if we set $f = (\lambda_1, \lambda_2, \mu, \nu)$, then $\langle f, y\rangle = 0$. Therefore $\operatorname{supp}\mu \subset \Omega_d$ and $\operatorname{supp}\nu \subset \Omega_u$. Now take $y \in Y$ positive. Then in particular $y_3 \geq 0$ on Ω_d and $y_4 \geq 0$ on Ω_u. Hence $y \in P$. By definition of $P^\perp$ this implies $f \geq 0$. In particular, $\mu, \nu \geq 0$. □

We now use the following lemma to characterize $\overline{V}^\perp$ in a different way.

LEMMA 5.9. *Let Y be a Banach space, and $P \subset Y$ a closed convex cone with dual cone $P^\perp$. Let X be a second Banach space, and $A : X \to Y$ a bounded linear operator. Let K be the following cone in X:*

$$K = \{u \in X : \ Au \in P\}.$$

Then the dual cone $K^\perp$ can be characterized by

$$K^\perp = \{A^T g \in X' : g \in P^\perp\}.$$

A proof of this lemma can be found in [11]. An immediate consequence of Lemma 5.9 is

COROLLARY 5.10.

$$\overline{V}^\perp = \{\beta^T(\lambda_1, \lambda_2, \mu, \nu) \in X' : (\lambda_1, \lambda_2, \mu, \nu) \in P^\perp\}.$$

We are now ready to derive the Euler-Lagrange equations.

PROOF OF THEOREM 5.3. We have seen that, since u is a minimizer, $H'(u) \in \overline{V}^\perp$ and

$$\overline{V} = \{v \in X \ : \ \beta v \geq 0\},$$

by Lemmas 5.6 and 5.7. By Corollary 5.10, there exists $(\lambda_1, \lambda_2, \mu, \nu) \in P^\perp$ such that $H'(u) = \beta^T(\lambda_1, \lambda_2, \mu, \nu)$. By Lemma 5.8, $\operatorname{supp}\mu \subset \Omega_d$,

$\operatorname{supp}\nu \subset \Omega_u$, and $\mu, \nu \geq 0$. The conjugate operator β^T is easily seen to be given by

$$\beta^T(\lambda_1, \lambda_2, \mu, \nu) = \lambda_1 - \lambda_2 + \mu\chi_{\Omega_d} - (\nu + \tau_{-h}\nu)\chi_{\Omega_u}, \tag{5.6}$$

Lastly, direct computation gives

$$H'(u) \cdot v = \int_{\mathbb{R}} \big[\ (\log(u/c_0 + 1) - 2\alpha\,\kappa * u)\, v\,\big]. \tag{5.7}$$

If we now redefine $\mu = \mu\chi_{\Omega_d}$, $\nu = \nu\chi_{\Omega_u}$, and $\lambda = \lambda_1 - \lambda_2$, we obtain the equation

$$\log(u/c_0 + 1) - 2\alpha\,\kappa * u = \lambda + \mu - \nu - \tau_{-h}\nu \quad \text{a.e.} \tag{5.8}$$

□

We still have to give the proof of Lemma 5.7. It uses the following result.

LEMMA 5.11. *Let $K \subset \mathbb{R}$ be a compact set. Let $V \subset \mathbb{R}$ be closed and set $V_\delta := V + (-\delta, \delta)$. Then*

$$|\{V_\delta \backslash V\} \cap K| \to 0 \quad \textit{as } \delta \to 0.$$

PROOF. Set $E = V \cap K$, and $E_\delta = V_\delta \cap K$. Then we can form the sequence

$$E_1 \supset E_{\frac{1}{2}} \supset \ldots \supset E_{\frac{1}{n}} \supset \ldots \supset F := \bigcap_{n=1}^{\infty} E_{\frac{1}{n}}.$$

Now we show that $E = F$. Evidently, $E \subset F$. Suppose that there exists an $x \in F \backslash E$. Then for any $n \in \mathbb{N}$, there exists an $x_n \in E$ such that $|x_n - x| \leq \frac{2}{n}$. But then x is an accumulation point of a sequence in E. Since E is closed, $x \in E$, a contradiction. Hence $E = F$. We now have (see, e.g., [144, p. 74]),

$$\lim_{n\to\infty} \left|E_{\frac{1}{n}}\right| = \left|\lim_{n\to\infty} E_{\frac{1}{n}}\right| = |E|.$$

This shows that

$$|E_\delta \backslash E| = |E_\delta| - |E| \to 0 \quad \text{as } \delta \to 0,$$

which proves the lemma. □

PROOF OF LEMMA 5.7. $\overline{V} \subset W$: Since $B : X \to Y$ is continuous, W is closed, and therefore it suffices to show that $V \subset W$. Take any $v \in V$. Then $B(u + \varepsilon_n v)(x) \geq 0$. Denote again the i-th component of a function u by u_i. Evidently, $b(v)_1 = b(v)_2 = 0$. Moreover, since

$$B(u + \varepsilon v) = B(u) + \varepsilon b(v),$$

$b(v)_3(x) \geq 0$ if $x \in \Omega_d$, $b(v)_4(x) \geq 0$ if $x \in \Omega_u$. Hence $\beta(v) \geq 0$, and $v \in W$.

$W \subset \overline{V}$: We prove the assertion for a dense subset in W, namely $W_c := \{w \in W \mid \operatorname{supp} w \text{ is compact}\}$. Then, since $\overline{V}$ is closed, $\overline{W}_c \subset \overline{V}$, i.e., $W \subset \overline{V}$.

First consider $w \in W_c$ such that $\operatorname{supp}(b(w)_4)_- \subset \Omega_u^c$, where $\operatorname{supp}(u)_-$ denotes the support of the negative part of u. Set $M_4 = \operatorname{supp}(b(w)_4)_-$. Then $M_4 \subset \operatorname{supp} w + \{0, h\}$, and hence M_4 is compact. Since $u \in H^1$ by Theorem 4.3, for $x \in \Omega_u$, $B(u)_4(x) = 0$, and since $\beta(w)_4(x) \geq 0$,

$$B(u + \varepsilon w)_4(x) \geq 0 \text{ for all } \varepsilon > 0 \text{ on } \Omega_u. \tag{5.9}$$

For the complement Ω_u^c, note that since M_4 is compact and $u \in H^1$ by Theorem 4.3, there exists a $\delta > 0$ such that

$$B(u)_4 \geq \delta > 0 \text{ on } M_4.$$

Hence, if we take ε_n such that

$$\varepsilon_n \leq N_u := \delta \left\| b(w)_4 \right\|_{L^\infty}^{-1},$$

then $B(u + \varepsilon_n w)_4 \geq 0$ on M_4.

Secondly, assume that $M_3 := \operatorname{supp}(b(w)_3)_- \subset \Omega_d^c$. For $x \in \Omega_d$, $B(u)_3(x) = 0$, and

$$B(u + \varepsilon w)_3(x) \geq 0 \text{ for all } \varepsilon > 0. \tag{5.10}$$

For $x \in \Omega_d^c$, in the same way as for $x \in \Omega_u^c$,

$$B(u + \varepsilon w)_3 \geq 0 \text{ for } \varepsilon \leq N_d := \delta \left\| b(w)_3 \right\|_{L^\infty}^{-1} \text{ on } M_3.$$

Hence, if both $\operatorname{supp}(b(w)_3)_- \subset \Omega_d^c$ and $\operatorname{supp}(b(w)_4)_- \subset \Omega_u^c$, and using that $B(u + \varepsilon w)_{1,2} = 0$ for any $\varepsilon > 0$, we find that for $\varepsilon \leq \min\{N_d, N_u\}$, $B(u + \varepsilon w) \geq 0$, and hence $w \in V$.

We now consider a general $w \in W_c$ and construct an approximation $\tilde{w}_\delta$ that satisfies the support assumptions above. Define for a given $\delta > 0$, $\Omega_{d,\delta} := \Omega_d + (-\delta, \delta)$, and $\Omega_{u,\delta} := \Omega_u + (-\delta, \delta)$. Define

$$w_\delta(x) = \begin{cases} 0 & \text{if } x \in \sum_{k \in \mathbf{Z}} \tau_{kh} S, \\ w(x) & \text{otherwise,} \end{cases}$$

where $S := (\Omega_{d,\delta} \backslash \Omega_d) \cup (\Omega_{u,\delta} \backslash \Omega_u)$. Since w has compact support, K is a finite set. The original w satisfies

$$w \geq 0 \text{ on } \Omega_d,$$
$$w + \tau_h w \leq w \text{ on } \Omega_u.$$

Hence w_δ satisfies

$$\begin{aligned} w_\delta &\geq 0 \text{ on } \Omega_{d,\delta}, \\ w_\delta + \tau_h w_\delta &\leq 0 \text{ on } \Omega_{u,\delta}, \end{aligned}$$

and therefore,

$$\begin{aligned} \operatorname{supp}(b(w_\delta)_3)_- &= \operatorname{supp}(w_\delta)_- \subset \Omega_{d,\delta}^c \subset \Omega_d^c, \\ \operatorname{supp}(b(w_\delta)_4)_- &= \operatorname{supp}(-w_\delta - \tau_h w_\delta)_- \subset \Omega_{u,\delta}^c \subset \Omega_u^c. \end{aligned}$$

Set $M := \int_{\mathbb{R}} w_\delta$. If $M = 0$, then set $\tilde{w}_\delta = w_\delta$.

If $M < 0$, then choose a smooth positive function ϕ, $\int \phi = 1$, with $\operatorname{supp}\phi \cap (\Omega_{u,\delta} + \{0, -h\}) = \emptyset$. Define $\tilde{w}_\delta := w_\delta - M\phi$ so that $\int \tilde{w}_\delta = 0$. In addition, $\tilde{w}_\delta \geq w_\delta \geq 0$ on $\Omega_{d,\delta}$, and $\tilde{w}_\delta + \tau_h \tilde{w}_\delta = w_\delta + \tau_h w_\delta \leq 0$ on $\Omega_{u,\delta}$.

If $M > 0$, choose a smooth positive function ϕ, $\int \phi = 1$, such that $\operatorname{supp}\phi \subset \mathbb{R}\backslash\Omega_{d,\delta}$. Define again $\tilde{w}_\delta := w_\delta - M\phi$ so that $\int \tilde{w}_\delta = 0$. Moreover, $\tilde{w}_\delta \geq 0$ on $\Omega_{d,\delta}$, and $\tilde{w}_\delta + \tau_h \tilde{w}_\delta \leq w_\delta + \tau_h w_\delta \leq 0$ on $\Omega_{u,\delta}$.

Therefore, in all cases, $\tilde{w}_\delta \in W_c$, such that

$$\begin{aligned} \operatorname{supp}(b(\tilde{w}_\delta)_3)_- &\subset \Omega_{d,\delta}^c \subset \Omega_d^c, \\ \operatorname{supp}(b(\tilde{w}_\delta)_4)_- &\subset \Omega_{u,\delta}^c \subset \Omega_u^c. \end{aligned}$$

Hence $\tilde{w}_\delta \in V$ by previous arguments. The proof is completed by showing that $\tilde{w}_\delta \to w$ in L^1. We have

$$\begin{aligned} \|\tilde{w}_\delta - w\|_{L^1} &= \|\hat{w}_\delta - M\phi - w\|_{L^1} \\ &\leq \|\hat{w}_\delta - w\|_{L^1} + M\|\phi\|_{L^1} \\ &\leq 2\|w\|_{L^\infty} \int_{\sum_{k\in\mathbf{Z}} \tau_{kh} S \cap \operatorname{supp} w} 1 \\ &\leq 2\|w\|_{L^\infty} \left| \sum_{k\in\mathbf{Z}} \tau_{kh} S \cap \operatorname{supp} w \right| \\ &\leq 2\|w\|_{L^\infty} \frac{\lceil |\operatorname{supp} w| + |S| \rceil}{h} |S|. \end{aligned}$$

By Lemma 5.11, in the limit as $\delta \to 0$,

$$|S \cap \operatorname{supp} w| = |(\Omega_{d,\delta}\backslash\Omega_d) \cup (\Omega_{u,\delta}\backslash\Omega_u)\} \cap \operatorname{supp} w| \to 0.$$

But then $\tilde{w}_\delta \to w$ in L^1 as $\delta \to 0$. □

5.3. Properties of $I(m)$

Recall that $I(m) = \inf\{H(u) \mid u \in K_m\}$. This section collects a number of results that elucidate some properties of $I(m)$. These will be used throughout this chapter.

LEMMA 5.12. *For any $m > 0$, $I(m) \geq -\frac{\alpha m^2}{2}$.*

PROOF. Since $(u, \kappa * u)_{L^2} \geq 0$ for any $u \in L^1(\mathbb{R})$,

$$\begin{aligned}(u, \kappa * u)_{L^2} &= |(u, \kappa * u)_{L^2}| \\ &\leq \|u\|_{L^1} \|\kappa * u\|_{L^\infty} \\ &\leq \|u\|_{L^1}^2 \|\kappa\|_{L^\infty} \\ &= \frac{1}{2} m^2.\end{aligned}$$

Therefore, using the positivity of $h(u)$,

$$H(u) \geq -\alpha (u, \kappa * u)_{L^2} \geq -\frac{\alpha m^2}{2}.$$

The right-hand side of this inequality is independent of u, and hence $I(m) \geq -\frac{\alpha m^2}{2}$. □

We now extend a well-known result from Lions [86, Thm. II.1] to our particular Problem (A). We will prove this theorem for Problem (B) in Section 5.4.

THEOREM 5.13. *Fix m, α, $h > 0$, $c_0 \in (0, 1/2)$. Every minimizing sequence for Problem (A) is compact if and only if*

$$I(m) < I(\mu) + I(m - \mu), \text{ for all } \mu \in (0, m).$$

PROOF. This is essentially Theorem II.1 from [86]. The only differences between that theorem and the current one are the conditions on $h(u)$ and the extra upper contact condition $u + \tau_h u \leq 1$. The addition of the upper contact condition does not interfere with the proof of Theorem II.1. It does, however, supply us with $\|u\|_{L^\infty} \leq 1$, which is a necessary condition, as discussed below. In [86], $h(u)$ is required to satisfy

$$h \text{ is strictly convex and non-negative,} \tag{5.11}$$

$$\lim_{t \to 0} h(t) t^{-1} = 0, \tag{5.12}$$

$$\lim_{t \to \infty} h(t) t^{-2} = \infty. \tag{5.13}$$

$h(u) = (u + c_0)\log(u/c_0 + 1) - u$ does satisfy (5.11) and (5.12), but not (5.13). However, (5.11)–(5.13) are required for any minimizing sequence $\{u_n\}$ in order to ensure

$$u_n \text{ is bounded in } L^1(\mathbb{R}) \cap L^2(\mathbb{R}), \tag{5.14}$$

$$h(u_n) \text{ is bounded in } L^1(\mathbb{R}). \tag{5.15}$$

Since

$$\|u\|_{L^2} \leq \|u\|_{L^1}^{1/2} \|u\|_{L^\infty}^{1/2} \leq m^{1/2},$$

(5.14) is easily satisfied for any sequence $\{u_n\} \subset K_m$. Now note that

$$\begin{aligned}\int u\, \kappa * u &\leq \|u\|_{L^2} \|\kappa * u\|_{L^2} \\ &\leq \|u\|_{L^2}^2 \|\kappa\|_{L^1} \\ &\leq \|u\|_{L^1} \|u\|_{L^\infty} \\ &\leq m.\end{aligned}$$

For any minimizing sequence $\{u_n\}$ we may assume, that $H(u_n) \leq C$ for some constant C. Hence, since $h(u)$ is strictly positive for any $u \in K_m$,

$$\|h(u)\|_{L^1} = \int h(u_n) = H(u_n) + \alpha \int u\, \kappa * u \leq C + \alpha m,$$

and $h(u_n)$ is bounded in $L^1(\mathbb{R})$. This means that (5.14) and (5.15) both apply. The rest of the proof is identical to the proof in [86]. □

The next two lemmas show two properties of $I(m)$ sketched in Figure 5.1 in the Introduction: $I(m)$ is always non-positive, and decreasing.

LEMMA 5.14. *For any $m > 0$, $I(m) \leq 0$.*

PROOF. We in fact prove a stronger result we will use later on: if $\{u_n\} \subset K_m$ is a minimizing sequence for Problem (A) that vanishes in the context of the Concentration-Compactness principle, then $\lim_{n\to\infty} H(u_n) = 0$.

Observe first that if $v \in K_m$, and setting $v_n = 1/n\, v(x/n)$, then $\lim_{n\to\infty} H(v_n) = 0$ (from which the lemma already immediately follows). It is therefore sufficient to prove that if $\{u_n\} \subset K_m$ is a vanishing minimizing sequence for Problem (A), then $\lim_{n\to\infty} H(u_n) \geq 0$.

Vanishing of u_n is equivalent to

$$\lim_{n\to\infty} \sup_{y\in\mathbb{R}} \int_{y-R}^{y+R} u_n = 0 \quad \text{for all } R < \infty. \tag{5.16}$$

We claim that $\|\kappa * u_n\|_{L^\infty} \to 0$ as $n \to \infty$. Pending the proof of this claim,

$$\int u_n \, \kappa * u_n = \int |u_n \, \kappa * u_n| \leq \|u_n\|_{L^1} \, \|\kappa * u_n\|_{L^\infty} \to 0 \text{ as } n \to \infty.$$

Since $h(u)$ is positive, it then follows directly that $\lim_{n\to\infty} H(u_n) \geq 0$.

To prove that $\|\kappa * u_n\|_{L^\infty} \to 0$, take u_{n_1} and u_{n_2} such that $u_n = u_{n_1} + u_{n_2}$ and such that for any $x \in \mathbb{R}$,

$$\begin{aligned} \operatorname{supp} u_{n_1} &\subset V_1 := (x - L, x + L), \\ \operatorname{supp} u_{n_2} &\subset V_2 := (-\infty, x - L] \cup [x + L, \infty). \end{aligned}$$

Using the positivity of u_n and κ,

$$\begin{aligned} \|\kappa * u_n\|_{L^\infty(V_1)} &= \sup_{x \in V_1} \int_{\mathbb{R}} \kappa(x - y) u_n(y) dy \\ &= \sup_{x \in V_1} \Big[\int_{x-L}^{x+L} \kappa(x - y) u_{n_1}(y) dy \\ &\qquad + \Big\{ \int_{-\infty}^{x-L} + \int_{x+L}^{\infty} \Big\} \kappa(x - y) u_{n_2}(y) dy \Big] \\ &\leq \sup_{y \in \mathbb{R}} \kappa(y) \cdot \sup_{x \in V_1} \int_{x-L}^{x+L} u_n(y) dy \\ &\qquad + \sup_{x \in V_1, y \in V_2} \kappa(x - y) \int_{\mathbb{R}} u_n(y) dy \\ &= \frac{1}{2} f_n(L) + m \frac{e^{-L}}{2}. \end{aligned}$$

Here, since u_n is a vanishing sequence,

$$f_n(L) := \sup_{x \in V_1} \int_{x-L}^{x+L} u_n \to 0 \text{ for any } L < \infty, \text{ as } n \to \infty.$$

Now for a given $\epsilon > 0$, choose L^ϵ large enough such that

$$\frac{1}{2} e^{-L} < \frac{\epsilon}{2},$$

and, for this L^ϵ, choose N_{L^ϵ} large enough such that

$$f_n(L^\epsilon) < \epsilon \text{ for all } n \geq N_{L^\epsilon}.$$

Then for all $n \geq N_{L^\epsilon}$

$$\|\kappa * u_n\|_{L^\infty(-L^\epsilon, L^\epsilon)} \leq \frac{1}{2} f_n(L^\epsilon) + m \frac{e^{-L^\epsilon}}{2} < \frac{\epsilon}{2}(1 + m).$$

We conclude $\|\kappa * u_n\|_{L^\infty} \to 0$ as $n \to \infty$. This completes the proof of the claim. □

REMARK 5.15. Lemma 5.14 also holds for Problem (B). The proof is identical.

LEMMA 5.16. *Let $m_2 > m_1 > 0$. Then $I(m_2) \leq I(m_1)$.*

PROOF. For a given $\varepsilon > 0$, choose $u \in K_{m_1}$ such that $H(u) \leq I(m_1) + \varepsilon$. Choose $v \in K_{m_2 - m_1}$ such that $u + v \in K_{m_2}$. Set $v_L(x) = 1/L\ v(x/L)$. Then also $u + \tau_L v_L \in K_{m_2}$ for L large enough, and, in addition,

$$H(u + \tau_L v_L) \leq I(m_1) + \varepsilon.$$

Therefore

$$I(m_2) \leq I(m_1) + \varepsilon.$$

Since $\varepsilon > 0$ was chosen arbitrarily, the lemma follows. □

In Lemma 5.30 in Section 5.5 we improve on this result by showing that if $2\alpha c_0 > 1$, then $I(m)$ strictly decreases with m.

Before we attempt to prove existence of minimizers for Problem (A), we first study a simplified problem, by dropping the upper contact condition $u + \tau_h u \leq 1$.

5.4. Dropping the upper contact condition

Recall the definition of Problem (B) on page 127. The following lemma will serve to make a bridge between Problems (A) and (B). It provides us with an L^∞-bound which scales linearly in mass m. This improves on the result in Remark 4.6 from Chapter 4, by using the regularity properties of minimizers. An immediate corollary of this lemma is that, if m is small enough, then minimizers of Problem (B) are also minimizers of Problem (A).

LEMMA 5.17. *For any minimizer u of Problem (B),*

$$\|u\|_{L^\infty} \leq \alpha(1 + c_0)\, \|u\|_{L^1}\,.$$

PROOF. Since u is continuous, on the support of u

$$\log(u/c_0 + 1) - 2\alpha\, \kappa * u = \lambda.$$

Hence

$$1 + u(x)/c_0 = e^{\lambda + 2\alpha\, \kappa * u(x)} \quad \text{for all } x \in \operatorname{supp} u.$$

Since κ' is bounded, $\kappa * u \in W^{1,\infty}$, so $u \in W^{1,\infty}$. Taking derivatives w.r.t. x yields

$$u' = 2\alpha(u + c_0)\kappa' * u \quad \text{on } \operatorname{supp} u.$$

Together with $u' = 0$ on $(\operatorname{supp} u)^c$, and using $u \leq 1$, this implies

$$\begin{aligned}
\left\|u'\right\|_{L^1(\mathbb{R})} &= 2\alpha \left\|(u + c_0)\kappa' * u\right\|_{L^1(\mathbb{R})} \\
&\leq 2\alpha \left\|u + c_0\right\|_{L^\infty(\mathbb{R})} \left\|\kappa' * u\right\|_{L^1(\mathbb{R})} \\
&\leq 2\alpha(1 + c_0) \left\|\kappa' * u\right\|_{L^1(\mathbb{R})} \\
&\leq 2\alpha(1 + c_0) \left\|\kappa'\right\|_{L^1(\mathbb{R})} \left\|u\right\|_{L^1(\mathbb{R})} \\
&= 2\alpha(1 + c_0) \left\|u\right\|_{L^1(\mathbb{R})}.
\end{aligned}$$

Now $\|u\|_{L^\infty} \leq \frac{1}{2} \|u'\|_{L^1}$, and therefore

$$\|u\|_{L^\infty} \leq \frac{1}{2} \|u\|_{L^1} \leq \alpha(1 + c_0) \|u\|_{L^1}.$$

□

The next corollary establishes the classical result from Lions [86, Thm. II.1]—already established for Problem (A) in Theorem 5.13—also for Problem (B).

COROLLARY 5.18. *Fix m, $\alpha > 0$, $0 < c_0 < 1/2$. Every minimizing sequence for Problem (B) is compact if and only if*

$$I(m) < I(\mu) + I(m - \mu), \text{ for all } \mu \in (0, m).$$

PROOF. By Lemma 5.17 we may choose minimizing sequences in the set

$$\left\{u \in K_m^B \mid \|u\|_{L^\infty} \leq \alpha(1 + c_0)m\right\}.$$

By the same reasoning as in Theorem 5.13, this implies that (5.14) and (5.15) hold and the method of proof of Theorem II.1 from [86] is applicable. □

The following lemma shows that, if we do not impose the upper contact condition, we can lower the energy of a function whose support consists of multiple disjoint sets by translating these sets towards each other.

LEMMA 5.19. *Let u be a minimizer of Problem (B). Then* $\operatorname{supp} u$ *is connected.*

PROOF. Suppose on the contrary that $u = u_1 + u_2$, such that $\operatorname{supp} u_1 = U_1$, $\operatorname{supp} u_2 = U_2$ and $U_1 \cap U_2 = \emptyset$. Without loss of generality we assume that if $x \in U_1$ and $y \in U_2$, then $x < y$. Let $u_l = \tau_l u_1 + u_2$. Then

$$\begin{aligned}
\frac{d}{dl}\left[\int_{\mathbb{R}} \tau_l u_1 \, \kappa * u_2 dx\right]_{l=0} &= \frac{d}{dl}\left[\int_{\mathbb{R}}\int_{\mathbb{R}} u_1(x-l)\kappa(x-y)u_2(y)dydx\right]_{l=0} \\
&= \frac{d}{dl}\left[\int_{\mathbb{R}}\int_{\mathbb{R}} u_1(x)\kappa(x+l-y)u_2(y)dydx\right]_{l=0} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} u_1(x)\kappa'(x-y)u_2(y)dydx \\
&= \int_{U_1}\left\{u_1(x)\int_{U_2}\kappa'(x-y)u_2(y)dy\right\}dx \\
&> 0
\end{aligned}$$

since $\kappa'(x-y) > 0$ for any $x \in U_1$ and $y \in U_2$. Then

$$\begin{aligned}
\frac{d}{dl}H(u_l)\Big|_{l=0} &= \frac{d}{dl}\Big[\int h(\tau_l u_1) + h(u_2) - \\
&\qquad \alpha[(\tau_l u_1)\,\kappa*(\tau_l u_1) + u_2\,\kappa*u_2 + 2(\tau_l u_1)\,\kappa*u_2]\Big]_{l=0} \\
&= -2\alpha\frac{d}{dl}\Big[\int (\tau_l u_1)\,\kappa*u_2\Big]_{l=0} \\
&< 0.
\end{aligned}$$

This contradicts with the assumption that u was a minimizer. □

REMARK 5.20. By exactly the same kind of reasoning, we can prove a similar lemma for minimizers of Problem (A). This states that if u is a minimizer, such that $u = u_1 + u_2$ and $\operatorname{supp} u_1 \cap \operatorname{supp} u_2 = \emptyset$, then $d(\operatorname{supp} u_1, \operatorname{supp} u_2) \leq h$. The proof is identical, and is based on the fact that we can still translate u_1 provided the two supports are more than a distance h apart.

The spherically decreasing rearrangement f^* of any nonnegative measurable function f on $\mathbb{R}^n$ is defined by

$$f^*(x) = \sup\{s > 0 \mid |N_s(f)| \geq \omega_n |x|^n\},$$

where

$$N_s(f) := \{x \in \mathbb{R}^n \mid f(x) > s\}$$

is the level set at height s, $|\cdot|$ denotes the Lebesgue measure, and ω_n is the measure of the unit ball in $\mathbb{R}^n$.

LEMMA 5.21 (Riesz, [112]). *The functional*

$$\mathcal{J}(f,g,h) := \int f(g * h)\, dx$$

does not decrease under spherical rearrangement. That is,

$$\mathcal{J}(f,g,h) \leq \mathcal{J}(f^*,g^*,h^*). \tag{5.17}$$

We restrict ourselves to functions u on $\mathbb{R}$, and hence call u^* the 'symmetric rearrangement' of u.

REMARK 5.22. From Riesz' Lemma 5.21 we can immediately deduce that if u is a minimizer for Problem (B), then so is u^*. We cannot yet conclude that $u = u^*$ though. Burchard has shown in [17] that if g is a fixed positive measurable symmetric strictly decreasing function, then a solution u that maximizes $\mathcal{J}(v,g,v)$ and its symmetric counterpart u^* are equal if and only if u^* is strictly symmetrically decreasing. Consequently, a minimizer u of Problem (A) and its symmetrically decreasing equivalent are equal if and only if u^* is strictly decreasing away from its maximum; such information is still lacking.

COROLLARY 5.23. *Let u be a minimizer of Problem (B). Then* $\operatorname{supp} u$ *is a bounded interval.*

PROOF. Note that by Lemma 5.21 u^* is also a minimizer. We claim that $\operatorname{supp} u^*$ is compact and hence an interval. Assume that $\operatorname{supp} u^*$ is not compact. Then $\operatorname{supp} u^* = \mathbb{R}$. Then by Theorem 8 in [11], u^* is an unstable stationary point, i.e., there exists $w \in \left\{ L^\infty(\mathbb{R}) \cap L^1(\mathbb{R}) \;\middle|\; \int w = 0 \right\}$ such that

$$H''(u^*) \cdot w \cdot w < 0,$$

which is a contradiction with the assumption that u^* is a minimizer. Finally, note that

$$|\operatorname{supp} u| = |\operatorname{supp} u^*|.$$

Lemma 5.19 now tells us that u has connected support, and hence $\operatorname{supp} u$ is a bounded interval. □

LEMMA 5.24. *Let u be a stationary point of Problem (B). Then $\lambda < 0$.*

PROOF. Suppose on the contrary that $\lambda \geq 0$. Since u is a stationary point, we have

$$\log(u/c_0 + 1) - 2\alpha\, \kappa * u = \mu + \lambda \geq \mu \geq 0.$$

By Corollary 5.23 u has compact support, and hence for x sufficiently large,

$$\mu \leq -2\alpha\, \kappa * u < 0.$$

This is a contradiction with the positivity of μ. □

5.5. Existence of minimizers: $2\alpha c_0 > 1$

We are now ready to prove the existence results for Problems (A) and (B). We distinguish two cases: $2\alpha c_0 > 1$ and $2\alpha c_0 < 1$. The difference between them is that in the first case for all $m > 0$ there exists a function $u \in K_m$ such that $H(u) < 0$. If $2\alpha c_0 < 1$ then we prove that for small m there are no admissible functions with negative energy. The significance of the first result is that these admissible functions with negative energy form the starting point of the proof of existence of minimizers for Problem (A).

Existence for Problem (B)

To prove existence of minimizers for Problem (A) for parameter values such that $2\alpha c_0 > 1$, we first have to derive that minimizers exist for Problem (B). This is done in a number of steps. We start by showing that for all $m > 0$ there exists a function $v \in L^2(\mathbb{R})$ with $\int v = 0$ such that

$$\frac{1}{c_0} \int v^2 - 2\alpha \int v\, \kappa * v < 0. \tag{5.18}$$

(Lemma 5.25). This lemma serves two purposes. It may be easily generalized to show that choosing parameters such that $2\alpha c_0 \geq 1$ is equivalent to the existence of admissible functions with negative energy for any $m > 0$ for both minimization problems (Lemma 5.26). This forms the starting point for a proof for existence of minimizers for Problem (B) given in Theorem 5.33. In addition, the lemma allows us to conclude that minimizers of Problem (A) have compact support (Lemma 5.29).

Recall that for any regularization kernel $\{\phi_n\}$ and any $u \in L^p(\mathbb{R})$, $\|\phi_n * u - u\|_{L^p} \to 0$ as $n \to \infty$ (see, e.g., [14, p. 71]).

LEMMA 5.25. *Suppose $2\alpha c_0 > 1$. Then there exists a $v \in L^2(\mathbb{R})$ with compact support and $\int v = 0$, such that*

$$\frac{1}{c_0} \int v^2 - 2\alpha \int v\, \kappa * v < 0. \tag{5.19}$$

PROOF. First we claim that if $2\alpha c_0 > 1$, then there exists a $v \in L^2(\mathbb{R})$, $\int v = 0$, such that (5.19) holds. It is sufficient to show that

$$\sup_{v \in L^2} \frac{\int v \, \kappa * v}{\int v^2} = 1.$$

Set $u_n = \frac{1}{n} u(\frac{x}{n})$, and $\kappa_n(x) = \kappa(nx)$. Then

$$\int u_n \, \kappa * u_n = \int u \kappa_n * u. \tag{5.20}$$

The sequence $\{n\kappa_n\}$ is a regularization kernel, and therefore

$$\begin{aligned} \left| \int u \, (n\kappa_n) * u - \int u^2 \right| &= \left| \int u[(n\kappa_n) * u - u] \right| \\ &\leq \|u\|_{L^2} \, \|(n\kappa_n) * u - u\|_{L^2} \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

This shows that

$$\frac{\int u_n^2}{\int u_n \, \kappa * u_n} = \frac{\frac{1}{n} \int u^2}{\frac{1}{n} \int u \, (n\kappa_n) * u} \to 1 \text{ as } n \to \infty.$$

So in fact, if we take any $w \in L^2$, and let $w_n = 1/n \, w(x/n)$, then (5.19) holds for n sufficiently large. We now show that such a w_n may be approximated by a function with compact support, while (5.19) remains valid.

The set $W = \{w \in C_c^\infty(\mathbb{R}) \mid \int w = 0\}$ lies dense in $\{w \in L^2(\mathbb{R}) \mid \int w = 0\}$. Take a sequence $v_n \subset W$, that converges strongly to v in the L^2-norm. Then

$$\begin{aligned} \left| \int v_n \, \kappa * v_n - \int v \kappa * v \right| &= \left| \int v_n \, \kappa * (v_n - v) - \int (v - v_n) \, \kappa * v \right| \\ &\leq \|v_n\|_{L^2} \, \|\kappa * (v_n - v)\|_{L^2} + \|v - v_n\|_{L^2} \, \|\kappa * v\|_{L^2} \\ &\leq \|v_n\|_{L^2} \, \|\kappa\|_{L^1} \, \|v_n - v\|_{L^2} + \|v - v_n\|_{L^2} \, \|\kappa * v\|_{L^2} \qquad (5.21) \\ &\to 0 \quad \text{as } n \to \infty. \end{aligned}$$

Hence, using (5.21)

$$\begin{aligned}\int v_n^2 - 2\alpha c_0 \int v_n \, \kappa * v_n &= \int [v_n^2 - v^2 + v^2] \\ &\quad - 2\alpha c_0 \int [v \, \kappa * v - v \, \kappa * v - v_n \, \kappa * v_n] \\ &\leq \int v^2 - 2\alpha c_0 \int v \, \kappa * v + o(1) \text{ as } n \to \infty \\ &< 0 \text{ for } n \text{ sufficiently large.}\end{aligned}$$

Hence for n large enough, (5.19) holds. This completes the proof. □

We now establish that if $2\alpha c_0 > 1$, there exists a function $u \in K_m$ with negative energy, and that if $2\alpha c_0 < 1$, then for m small enough all admissible functions have non-negative energy. This lemma may be extended into a true equivalence between $2\alpha c_0 \geq 1$ and existence of admissible functions for all $m > 0$ with negative energy.

LEMMA 5.26. *If $2\alpha c_0 > 1$, then for all $m > 0$ there exists a function $u \in K_m$ such that $H(u) < 0$. If $2\alpha c_0 < 1$, then for $m > 0$ small enough there does not exist any function in K_m such that $H(u) < 0$.*

PROOF. Assume first that $2\alpha c_0 > 1$. As may be seen from the proof in Lemma 5.25, for any $m > 0$ we can find a function u in $L^2(\mathbb{R})$ with $\int u = m$, such that

$$\frac{1}{2c_0} \int u^2 - \alpha \int u \, \kappa * u < 0.$$

In fact, from the construction of u in that same proof, it is also immediate that we may assume that $u \in K_m$. Since $h(u) < \frac{1}{2c_0} u^2$, we conclude that for all $m > 0$ there exists $u \in K_m$ such that $H(u) < 0$.

Now assume that $2\alpha c_0 < 1$. We show that m can be chosen small enough such that $H(u) \geq 0$ for all $u \in K_m$.

Define the convolution kernel $\kappa^{1/2}$ by

$$\kappa^{1/2} * \kappa^{1/2} * u = \kappa * u \quad \text{for all } u \in L^1(\mathbb{R}).$$

Then

$$\int u \, \kappa * u = \int (\kappa^{1/2} * u)^2.$$

Since h is a convex function of u, Jensen's inequality gives

$$\int \kappa^{1/2} * h(u) \geq \int h(\kappa^{1/2} * u). \tag{5.22}$$

In addition, $\int \kappa^{1/2} = \widehat{\kappa^{1/2}}(0) = 1$. Therefore, by Fubini, $\int h(u) = \int \kappa^{1/2} * h(u)$. Now note that $h''(x) = \frac{1}{x+c_0}$. Hence, setting $\delta := \left\|\kappa^{1/2} * u\right\|_{L^\infty}$,

$$h''(\kappa^{1/2} * u) \geq \frac{1}{\delta + c_0},$$

and

$$h(\kappa^{1/2} * u) \geq \frac{1}{2(\delta + c_0)} \int (\kappa^{1/2} * u)^2. \tag{5.23}$$

Using (5.22) and (5.23),

$$\begin{aligned} H(u) &= \int h(u) - \alpha \int (\kappa^{1/2} * u)^2 \\ &= \int \kappa^{1/2} * h(u) - \alpha \int (\kappa^{1/2} * u)^2 \\ &\geq \int h(\kappa^{1/2} * u) - \alpha \int (\kappa^{1/2} * u)^2 \\ &\geq \left(\frac{1}{2(\delta + c_0)} - \alpha \right) \int (\kappa^{1/2} * u)^2. \end{aligned}$$

Since $\delta = \left\|\kappa^{1/2} * u\right\|_{L^\infty} \leq \left\|\kappa^{1/2}\right\|_{L^\infty} \|u\|_{L^1} = \left\|\kappa^{1/2}\right\|_{L^\infty} m$, we may take m small enough, such that

$$\frac{1}{2(\delta + c_0)} - \alpha > 0. \tag{5.24}$$

But then, by positivity of $\int (\kappa^{1/2} * u)^2$, $H(u) \geq 0$. □

REMARK 5.27. Note that, since $\left\|\kappa^{1/2} * u\right\|_{L^\infty} \leq \|u\|_{L^\infty}$, if we replace $\delta = \left\|\kappa^{1/2} * u\right\|_{L^\infty}$ by $\|u\|_{L^\infty}$ in equation (5.24), we obtain a sufficient condition on u such that its energy is non-negative:

$$\text{If } \|u\|_{L^\infty} < \frac{1 - 2\alpha c_0}{2\alpha}, \text{ then } H(u) \geq 0.$$

If this is the case, then by Lemma 5.14 there are minimizing sequences that are not compact in the sense of the Concentration-Compactness principle. This may suggest that minimizers with small L^∞-norm do not exist. We can also reverse the argument and conclude that, since $\|u\|_{L^\infty} \leq 1$, there are no admissible functions with negative energy for any $m > 0$ if

$$\alpha \leq \frac{1}{2(1 + c_0)}, \tag{5.25}$$

or, equivalently, if

$$c_0 \leq \frac{1-2\alpha}{2\alpha}. \tag{5.26}$$

Note that this last inequality gives us a first hint towards explaining the Critical Micelle Concentration discussed in Chapter 1: if $\alpha \in (1/3, 1/2)$, then (5.26) provides a (non-trivial) upper bound on c_0 below which not all minimizing sequences are compact, suggesting that minimizers cease to exist. For $\alpha \leq 1/3$ this holds for all $c_0 \in (0, 1/2)$.

We will reflect on this when we present the numerical results in Section 5.8.

The first existence result deals with Problem (B). It will serve as the start of the existence of minimizers of Problem (A) in Theorem 5.32.

THEOREM 5.28. *Suppose* $2\alpha c_0 > 1$. *Then for all* $m > 0$, *there exists a minimizer for Problem (B).*

PROOF. By Lemma 5.25 we know that for any $m > 0$ there exists a $v \in L^2(\mathbb{R})$ such that

$$\frac{1}{2c_0}\int v^2 - \alpha \int v\, \kappa * v < 0. \tag{5.27}$$

In fact, it follows from the proof of Lemma 5.25 that if we take any $v \in L^2$, and construct the sequence $v_n = 1/n\, v(x/n)$, then (5.27) holds for n sufficiently large. Note also that if $v \in K_m^B$, then $v_n \in K_m^B$ for all $n \in \mathbb{N}$. Let us therefore take a function $u \in K_m^B \cap L^2(\mathbb{R})$ such that (5.27) is satisfied. Write $H(u) = \int [h(u) - \alpha u\, \kappa * u]$ as before. Then, using Taylor series it is easy to see that $h(x) < 1/(2c_0)x^2$ for all $x > 0$. Hence if v satisfies (5.27), then also $H(u_m) < 0$. Hence

$$I(m) \leq H(u_m) < 0 \text{ for any } m > 0.$$

We now show that $I(m)/m$ is a strictly decreasing function for all $m > 0$. First we claim that

$$H'(u) \cdot u \leq 2H(u). \tag{5.28}$$

Postponing the proof of this inequality for the moment, we proceed as follows. Let $\{u_n\} \subset K_m^B$ be a minimizing sequence. Then

$$\frac{d}{d\mu}\frac{H(\mu u_n)}{\mu m}\Big|_{\mu=1} = \frac{1}{m}\left(H'(u_n)\cdot u_n - H(u_n)\right) \leq \frac{H(u_n)}{m}.$$

We may assume that the minimizing sequence is chosen such that $H(u_n) \leq I(m)/2$. Then

$$\limsup_{n\to\infty} \frac{d}{d\mu}\frac{H(\mu u_n)}{\mu m}\Big|_{\mu=1} \leq \frac{I(m)}{2} < 0.$$

We infer $I(m)/m$ is a strictly decreasing function for all $m > 0$. This is a sufficient condition to ensure compactness of any minimizing sequence: for any $\theta \in (0,1)$,

$$\begin{aligned} I(m) = m\frac{I(m)}{m} &< m\left(\theta\frac{I(\theta m)}{\theta m} + (1-\theta)\frac{I((1-\theta)m)}{(1-\theta)m}\right) \\ &= I(\theta m) + I((1-\theta)m). \end{aligned} \tag{5.29}$$

Theorem 5.13 now ensures compactness of any minimizing sequence. Minimizers for Problem (B) now exist, since the limit of a minimizing sequence is in fact a minimizer [86].

We are left with proving (5.28). Let $H(u) = \int [h(u) - \alpha u\, \kappa * u]$ as before. Then

$$\begin{aligned} \frac{1}{c_0}\left(H'(u)\cdot u - 2H(u)\right) &= \frac{1}{c_0}\int \left[uh'(u) - 2h(u)\right] \\ &= \frac{u}{c_0}\int \Big[\log\left(1+\frac{u}{c_0}\right) \\ &\qquad - 2\left(1+\frac{u}{c_0}\right)\log\left(1+\frac{u}{c_0}\right) - 2\frac{u}{c_0}\Big]. \end{aligned}$$

Setting $x = u/c_0$, and rearranging terms, it remains to be shown that

$$\log(1+x) \geq \frac{2x}{2+x} \qquad \text{for} \quad x \geq 0. \tag{5.30}$$

Note that in $x = 0$, the left and right terms are equal. For any $x > 0$,

$$\frac{d}{dx}\left(\log(1+x) - \frac{2x}{2+x}\right) = \frac{x^2}{(1+x)(2+x)^2} > 0,$$

and therefore (5.30) indeed holds. This completes the proof. □

A different method of proof for existence of minimizers of Problem (B) is presented in the Appendix. It shows that in the particular case of Problem (B), any minimizing sequence that splits in the sense of the Concentration-Compactness Lemma is in fact compact in L^1, but the limit function has lower mass.

Existence for Problem (A)

The proof for existence of minimizers for Problem (A), given in Theorem 5.32 uses a number of ingredients.

We start by showing that if $2\alpha c_0 > 1$, then minimizers have compact support (Lemma 5.29). This is an essential ingredient in the proof

for existence of minimizers later on, since it allows one to add two minimizers after a proper translation—thereby constructing an admissible function with greater mass—while keeping control over the energy.

Then we show that $I(m)$ decreases sufficiently fast with m in Lemma 5.30.

LEMMA 5.29. *Suppose $2\alpha c_0 > 1$. Let u be a minimizer of Problem (A). Then $\lambda < 0$, and u has compact support.*

PROOF. Assume that $\lambda \geq 0$. By Lemma 5.2, Ω_u is compact. Since u is a stationary point, we have for $x \in \Omega_u^c$,

$$\log(u(x)/c_0 + 1) = 2\alpha\, \kappa * u(x) + \mu + \lambda > 0,$$

and hence $u > 0$ on Ω_u^c. By Lemma 5.25 there exists a $v \in L^2(R)$ with compact support and $\int v = 0$, such that

$$\frac{1}{c_0} \int v^2 - 2\alpha \int v\, \kappa * v < 0.$$

Since v has compact support, we can translate v such that

$$\operatorname{supp} \tau_{-L} v \subset\subset [L, \infty) \subset \operatorname{supp} u,$$

for a suitable $L \in \mathbb{R}$. Set $v_L := \tau_{-L} v$. Assume L is large enough such that $\Omega_u \cap [L, \infty) = \emptyset$. By Theorem 4.3, u is continuous. Hence there exists $\bar{u} \in \mathbb{R}$ such that $u \geq \bar{u} > 0$ on $\operatorname{supp} v_L$. Therefore

$$u + \varepsilon v_L \geq 0 \text{ if } \varepsilon \leq \frac{\|v_L\|_{L^\infty}}{\bar{u}}.$$

Moreover, setting $B_L = \operatorname{supp} v_L$, $\|u\|_{L^\infty(B_L)} \to 0$ as $L \to \infty$. Hence,

$$\begin{aligned} \|u + \varepsilon v_L\|_{L^\infty(B_L)} &\leq \|u\|_{L^\infty(B_L)} + \varepsilon \|v_L\|_{L^\infty(B_L)} \\ &< \frac{1}{2} \text{ if } \varepsilon < \frac{\frac{1}{2} - \|u\|_{L^\infty(B_L)}}{\|v_L\|_{L^\infty(B_L)}}. \end{aligned}$$

If we take L large enough such that $\|u\|_{L^\infty(B_L)} < \frac{1}{2}$, then there exists $\varepsilon > 0$ such that $u + \varepsilon v_L \in K_m$. Hence, in that case v_L is an unstable direction for u:

$$\begin{aligned} H''(u) \cdot v_L \cdot v_L &= \int \frac{1}{u + c_0} (v_L)^2 - 2\alpha \int v_L\, \kappa * v_L \\ &< \int \frac{1}{c_0} (v_L)^2 - 2\alpha \int v_L\, \kappa * v_L \\ &< 0. \end{aligned}$$

This is a contradiction with the assumption that u is a minimizer.

So we know that $\lambda < 0$. Suppose that u does not have compact support. By Theorem 4.3, u is continuous. Take a sequence $\{x_n\}$, $x_n \to \infty$, such that $u(x_n) > 0$. Then for n large enough, $x_n \notin \Omega_u + \{0, h\}$, since Ω_u is compact. For each n there exists an $\varepsilon_n > 0$ such that u is strictly positive on $(x_n - \varepsilon_n, x_n + \varepsilon_n)$. Hence, for n sufficiently large, $\mu = 0$ on $(x_n - \varepsilon_n, x_n + \varepsilon_n)$. Therefore, for $\bar{x}_n \in (x_n - \varepsilon_n, x_n + \varepsilon_n)$,

$$\lambda + 2\alpha(\kappa * u)(\bar{x}_n) = \log(1 + u(\bar{x}_n)/c_0) > 0.$$

In the limit as $n \to \infty$, $(\kappa * u)(\bar{x}_n) \to 0$ and hence the left-hand side is negative, a contradiction. □

LEMMA 5.30. *Assume $2\alpha c_0 > 1$. Suppose there exists a minimizer $u \in K_m$ for some m for Problem (A). Then there exists a constant $c > 0$ such that for $\varepsilon > 0$ sufficiently small*

$$I(m + \varepsilon) < I(m) - c\varepsilon.$$

PROOF. By Lemma 5.29, u has compact support. Note that for $u \in K_m$, and for any $v \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$,

$$\begin{aligned} H''(u) \cdot v \cdot v &= \int \frac{v^2}{u + c_0} - 2\alpha \int v \, \kappa * v \\ &\leq \frac{1}{c_0} \int v^2 \\ &\leq \frac{1}{c_0} \|v\|_{L^1} \|v\|_{L^\infty}. \end{aligned}$$

Hence, using $H(u + \varepsilon v) = H(u) + \varepsilon H'(u) \cdot v + \varepsilon^2 H''(\theta u) \cdot v \cdot v$ for a suitable $\theta \in [0, 1]$, we know that

$$H(u + \varepsilon v) - H(u) \leq \varepsilon H'(u) \cdot v + O(\varepsilon^2 \|v\|_{L^1} \|v\|_{L^\infty}). \tag{5.31}$$

Take a positive function $v \in L^1 \cap L^\infty$, $\int v = 1$, with support an interval of length 1, and $\|v\|_{L^\infty} = 1$. Consider $w = u + \varepsilon v$, where v is translated such that

$$d(\operatorname{supp} u, \operatorname{supp} v) > h. \tag{5.32}$$

Then $w \in K_{m+\varepsilon}$ for ε sufficiently small. Observe that there exist a positive constant C such that

$$\|\kappa * u\|_{L^\infty} \geq C > 0 \text{ on } \operatorname{supp} v.$$

We therefore have

$$H'(u) \cdot v = -2\alpha \int_{\mathbb{R}} v \, \kappa * u < -2\alpha C < 0 \text{ on } \operatorname{supp} v. \tag{5.33}$$

By translating v such that $d(\operatorname{supp} u, \operatorname{supp} v)$ decreases, we improve on $H(u+\varepsilon v)$, while ensuring that the new function still lies in $K_{m+\varepsilon}$ by (5.32) (see Lemma 5.19). Hence by (5.31) and (5.33), we obtain for ε sufficiently small,

$$I(m+\varepsilon) < H(u+\varepsilon v) \leq H(u) - 2\varepsilon\alpha C.$$

This completes the proof. □

REMARK 5.31. Lemma 5.30 also holds for Problem (B).

THEOREM 5.32. *Fix $h > 0$ and assume that $2\alpha c_0 > 1$. Then for any $m > 0$ there exists a minimizer of Problem (A).*

PROOF. The proof follows that of Theorem 2.2 in [108]. Let $I(m)$ be the minimal value of H for a given $m > 0$. By Theorem 5.13 it is sufficient to prove that

$$I(m_1 + m_2) < I(m_1) + I(m_2) \text{ for all } m_1, m_2 > 0.$$

Denote

$$A := \bigcup \{(0, \bar{m}] \mid \text{For all } 0 < m < \bar{m} \text{ Problem (A) has a minimizer}\}.$$

We have to prove that $A = \mathbb{R}^+$. We prove first that A is non-empty. Then we show A is closed, and lastly that A is also open: $\mathbb{R}^+$ is the only set within $\mathbb{R}^+$ that is both open and closed, and not empty.

A **is non-empty:** Let for a given $m > 0$, u be a minimizer of Problem (B) by Theorem 5.28. Recall that by Lemma 5.17,

$$\|u\|_{L^\infty} \leq \alpha m(1 + c_0).$$

Hence, for m sufficiently small, $\|u\|_{L^\infty} \leq 1/2$, and $u \in K_m$. This implies u is then a minimizer for Problem (A).

A **is closed:** By Lemma 5.29, u has compact support. Take m_1 and m_2 sufficiently small to ensure the existence of minimizers u_1 and u_2 respectively, and consider a new function $v := u_1 + u_2$, where u_1 and u_2 are translated if necessary to ensure they have disjunct supports. Then as in (5.41),

$$H(v) - H(u_1) - H(u_2) < 0.$$

This yields

$$I(m_1 + m_2) \leq H(v) < H(u_1) + H(u_2) = I(m_1) + I(m_2).$$

This argument shows that A is closed in $\mathbb{R}^+$: if $(0, \bar{m}) \subset A$ then for any $\theta \in (0, 1)$, we have existence of minimizers for mass $\theta\bar{m}$. By the above

argument

$$I(\bar{m}) \leq H(u_{\bar{m}}) < H(u_{\theta\bar{m}}) + H(u_{(1-\theta)\bar{m}}) = I(\theta\bar{m}) + I((1-\theta)\bar{m}).$$

Hence $(0, \bar{m}] \subset A$ by Theorem 5.13.

A **is open:** Suppose that there exists an $m_1 > m$ such that $m_1 \notin A$. Then there exists an $m_2 < m_1$, such that

$$I(m_1) = I(m_2) + I(m_1 - m_2). \tag{5.34}$$

If $m_2, m_1 - m_2 \in A$, then there exist $u \in K_{m_2}$ and $v \in K_{m_1 - m_2}$ such that $I(m_2) = H(u)$ and $I(m_1 - m_2) = H(v)$. But then by previous arguments, we can construct a function $w \in K_{m_1}$ by summing u and v after appropriate translations, and we find

$$I(m_1) \leq H(w) < H(u) + H(v) = I(m_2) + I(m_1 - m_2).$$

This is in contradiction with (5.34). So we may assume without loss of generality that $m_2 \notin A$.

Applying this argument inductively, we find a sequence $\{m_k\}_{k=1}^{\infty} \subset A^c$ such that $m_k > m_{k+1}$ for all $k \in \mathbb{N}$. This sequence is bounded and hence has a limit, which we will denote by m_1^∞. If $m_1^\infty \notin A$, then we may repeat the construction, and find a new limit $m_2^\infty < m_1^\infty$, etcetera. The sequence $\{m_n^\infty\}$ we thus construct converges to $\bar{m} = \sup A$. To simplify notation, we drop the ∞ superscripts.

Now for any $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $m_n - \bar{m} < \varepsilon$ for all $n \geq N$. By Lemma 5.12,

$$I(\varepsilon) \geq -\frac{\alpha}{2}\varepsilon^2.$$

Therefore,

$$I(m_n) - I(m_{n+1}) = I(m_{n+1} - m_n) \geq -\frac{\alpha}{2}(m_n - m_{n+1})^2,$$

and

$$\begin{aligned} I(m_N) - I(\bar{m}) &\geq -\frac{\alpha}{2} \sum_{n \geq N} (m_n - m_{n+1})^2 \\ &\geq -\frac{\alpha}{2}(\bar{m} - m_N)^2 \qquad (5.35) \\ &\geq -\frac{\alpha}{2}\varepsilon^2. \qquad (5.36) \end{aligned}$$

But by Lemma 5.30, there exists a constant $c > 0$ such that

$$I(m_N) - I(\bar{m}) < -c(\bar{m} - m_N) < -c\varepsilon.$$

Taking the limit $\varepsilon \to 0$ we obtain a contradiction with (5.35). This proves that A is open. □

5.6. Existence of minimizers: $2\alpha c_0 < 1$

For values of α and c_0 such that $2\alpha c_0 < 1$, we do not expect that we have existence for all $m > 0$ in the light of Lemma 5.26. However, for Problem (B) we can prove that there exist minimizers if m is sufficiently large. We have seen in Lemma 5.14 that vanishing minimizing sequences have zero energy in the limit. The main argument to prove existence then, is that if for a given m there exists a function $u \in K_m^B$ such that $H(u) < 0$, then we have existence of minimizers.

THEOREM 5.33. *Fix $\alpha > 0$, $c_0 \in (0, 1/2)$. There exists an $\bar{m} > 0$ such that for all $m > \bar{m}$ Problem (B) has a minimizer.*

PROOF. Let $\{u_n\}$ be a minimizing sequence. If the sequence vanishes, then by applying the argument in the proof of Lemma 5.14 we find that

$$\lim_{n\to\infty} H(u_n) = 0.$$

We construct a function u such that for m sufficiently large,

$$H(u) < 0.$$

This proves that the minimizing sequence cannot vanish for m large enough. Set

$$u(x) = \begin{cases} M & \text{for } 0 < x < m/M, \\ 0 & \text{otherwise.} \end{cases}$$

Using the explicit calculation

$$\int \chi_{[0,\ell]} \kappa * \chi_{[0,\ell]} = e^{-\ell} - 1 + \ell,$$

where χ_A is the characteristic function of the set A, we compute

$$\begin{aligned} H(u) &= \frac{m}{M}(M + c_0)\log(M/c_0 + 1) - m - \alpha \int u\, \kappa * u \\ &= \frac{m}{M}(M + c_0)\log(M/c_0 + 1) - m - \alpha M^2 \left[e^{-\frac{m}{M}} - 1 + \frac{m}{M}\right]. \end{aligned}$$

For M large enough, there exists $\bar{m}$ such that for all $m > \bar{m}$, $H(u) < 0$. Hence, for such m we conclude that $\{u_n\}$ cannot vanish.

Note that by the same proof as in Theorem 5.28, $I(m)/m$ is a strictly decreasing function for all $m > 0$ such that $I(m) < 0$. Therefore, by the same reasoning as in Theorem 5.28, for all $m > \bar{m}$,

$$I(m) < I(\mu) + I(m - \mu) \text{ for all } \mu \in (0, m),$$

and existence of minimizers may be concluded. □

5.7. Results on the structure of Ω_u

In this section we collect a few results on the upper contact set Ω_u, chiefly concerning its size (Lemmas 5.35 and 5.36), and the periodicity of minimizers on Ω_u (Lemma 5.37). We also state a conjecture on the structure of Ω_u and the intersection of Ω_u and Ω_d (Conjecture 5.1).

LEMMA 5.34. *Assume u is a stationary point of Problem (A), and that u has neither upper or lower contact on an open interval I (possibly unbounded). Then*

$$\left\|u'\right\|_{L^\infty} \leq 2\alpha(1+c_0) \text{ on } I.$$

PROOF. As in Lemma 5.17,

$$u' = 2\alpha(u+c_0)\,\kappa' * u \text{ a.e. on } I.$$

Hence, using $u \leq 1$,

$$\begin{aligned}\left\|u'\right\|_{L^\infty(I)} &= \left\|2\alpha(u+c_0)\,\kappa' * u\right\|_{L^\infty(I)} \\ &\leq 2\alpha(1+c_0)\left\|\kappa' * u\right\|_{L^\infty(I)} \\ &\leq 2\alpha(1+c_0)\left\|\kappa'\right\|_{L^1(I)}\|u\|_{L^\infty(I)} \\ &\leq 2\alpha(1+c_0). \end{aligned} \tag{5.37}$$

□

LEMMA 5.35. *Let $u \in K_m$, and let Ω_u be the upper contact set. Suppose that $\Omega_u = [0, L]$, and assume $L > h$. Then*

$$L \leq 2\|u\|_{L^1} - h + (L+h) \bmod 2h.$$

PROOF. By Lemma 5.37, $u(x) = u(x+2h)$ for all $x \in [0, L-h]$. Denote $N = \lfloor \frac{L+h}{2h} \rfloor$. Then

$$\begin{aligned}\int_0^{L+h} u &= N\int_0^{2h} u + \int_{2hN}^{L+h} u \\ &= N\left(\int_0^h u + \int_h^{2h} u\right) + \int_{2hN}^{L+h} u \\ &= N\left(\int_0^h u + \int_0^h 1-u\right) + \int_{2hN}^{L+h} u \\ &= Nh + \int_{2hN}^{L+h} u \\ &\geq Nh.\end{aligned}$$

Therefore

$$\begin{aligned} L+h &= 2hN + (L+h) \bmod 2h \\ &\leq 2\int_0^{L+h} u + (L+h) \bmod 2h \\ &\leq 2\|u\|_{L^1} + (L+h) \bmod 2h. \end{aligned}$$

□

More generally,

LEMMA 5.36. *Let $u \in K_m$, and let Ω_u be the upper contact set. Then Ω_u has finite measure. In fact,*

$$\mu(\Omega_u) \leq 2\|u\|_{L^1}.$$

If $\Omega_u \cap (\Omega_u + h) = \emptyset$, then

$$\mu(\Omega_u) \leq \|u\|_{L^1}.$$

PROOF.

$$\mu(\Omega_u) = \int_{\Omega_u} [u + 1 - u] = \int_{\Omega_u} u + \int_{\Omega_u + h} u.$$

Hence, a priori, $\mu(\Omega_u) \leq 2\|u\|_{L^1}$. If $\Omega_u \cap (\Omega_u + h) = \emptyset$, then

$$\int_{\Omega_u} u + \int_{\Omega_u + h} u \leq \|u\|_{L^1}.$$

□

LEMMA 5.37. *Let u be a stationary point of Problem (A). Suppose $[x_0, x_1] \subset \Omega_u$, $x_1 - x_0 > h$. Then $u - \tau_{2h}u = 0$ on $[x_0, x_1 - h]$.*

PROOF. On $[x_0, x_1]$, $u = 1 - \tau_h u$. If $x \in [x_0, x_1 - h]$, then $x + h \in [x_0 + h, x_1] \subset \Omega_u$. Hence on that interval $\tau_h u = 1 - \tau_{2h}u$. Combining these two equalities we obtain the desired result. □

COROLLARY 5.38. *Let u be a stationary point of Problem (A). Suppose $[x_0, x_1] \subset \Omega_u$, $x_1 - x_0 > h$.*

- *If $\Omega_u \cap \Omega_d = \emptyset$, then $\nu + \tau_{-h}\nu = \tau_{2h}(\nu + \tau_{-h}\nu)$ on $[x_0, x_1 - h]$.*
- *If $\Omega_u \cap \Omega_d \neq \emptyset$, then $\mu - \nu - \tau_{-h}\nu = \tau_{2h}(\mu - \nu - \tau_{-h}\nu)$ on $[x_0, x_1 - h]$.*

5.8. Numerical simulations

We have performed some simulations to investigate the minimizers of Problem (A) using an L^2-gradient flow, programmed in Matlab. Solutions are computed on a domain $[-L, L]$, discretized using uniform grids. To compute solutions that conform to the contact constraints

$$u \geq 0, \quad u + \tau_h u \leq 1,$$

we introduce energy penalization terms,

$$I_1 = C_1 \int_{-L}^{L} (\min\{u, 0\})^2, \quad I_2 = C_2 \int_{-L}^{L} (\max\{u + \tau_h u - 1, 0\})^2.$$

The resulting functional

$$\int_{-L}^{L} [(u + c_0) \log(u/c_0 + 1) - \alpha u \, \kappa * u] + I_1 + I_2$$

then has the following Euler-Lagrange equation

$$f(u) := \log(u/c_0 + 1) - 2\alpha \, \kappa * u + I_1' + I_2' = 0, \tag{5.38}$$

where

$$I_1' = 2C_1 \min\{u, 0\},$$

and

$$I_2' = -2C_2 \left[\max\{u + \tau_h u - 1, 0\} + \max\{\tau_{-h} u + u - 1, 0\}\right].$$

The gradient flow is generated by

$$u_t = -f(u) - \frac{1}{2L} \int_{-L}^{L} f(u), \tag{5.39}$$

where the last term maintains the third constraint of Problem (A), mass conservation during the flow. Indeed, if we set $u(t, x)$ to be the computed solution at time t, then

$$\begin{aligned} \frac{d}{dt} \int_{-L}^{L} u(t, x) dx &= \int_{-L}^{L} u_t(t, x) dx \\ &= \int_{-L}^{L} \left\{ -f(u(t, x)) - \frac{1}{2L} \int_{-L}^{L} f(u(t, x)) \right\} dx \\ &= 0. \end{aligned}$$

Note that for some stationary points, $\Omega_u \cap \Omega_d \neq \emptyset$ (at least in the simulations). This requires us to set $C_1 = C_2$.

The convolution $\kappa * u$ is computed using the trapezoid rule. The gradient flow defined by (5.39) is solved using the `ode23` solver from

Matlab. For $2\alpha c_0 > 1$, minimizers of Problem (A) have compact support by Lemma 5.29. We suspect this to be true for $2\alpha c_0 < 1$ as well, and assume we can always extend the solution outside $[-L, L]$ by 0. Then there are no boundary effects that have to be taken into account if $\operatorname{supp} u \subset [-L, L]$, since then for any $x \in [-L, L]$,

$$\int_{\mathbb{R}} \kappa(x-y)u(y)dy = \int_{-L}^{L} \kappa(x-y)u(y)dy,$$

which is exactly what we compute.

Note that we do not expect the gradient flow to always end up at a stationary point in actual computations: when we take an initial condition consisting of two compactly supported parts that are a great distance apart, then the attraction force from the convolution integral is only exponentially small in the distance between the two parts. These two lumps of mass will stay put during the gradient flow when the force is dominated by the numerical noise in the computation. This indicates that initial conditions have to be chosen with care. We will return to this later on.

Existence of minimizers

We have discussed in detail the existence of minimizers for $2\alpha c_0 > 1$ in previous sections, culminating in Theorem 5.32. Here we explore existence of solutions for $2\alpha c_0 < 1$, using the gradient flow setup outlined above.

For Problem (B) we have proved in Theorem 5.33 that minimizers exist for m sufficiently large. Lemma 5.26 moreover states that if $2\alpha c_0 < 1$, then there are no admissible functions with negative energy for small enough m. These two results indicate that minimizers may cease to exist for small m, and this is explored in the simulations.

In Figure 5.3 the existence of non-trivial solutions to Problem (A) in the (m, α)-plane is depicted. The figure is computed on a 100×100 grid, performing continuation in decreasing m: for each α, an initial condition is chosen at $m = 6$, which converges to a non-trivial stationary point; then, this stationary point is rescaled to smaller mass by simple multiplication of the function by an appropriate constant, and the gradient flow is continued.

At each grid point, the gradient flow is computed for a particular fixed time span, each time continuing from the last computed solution. Let $u(\mathbf{x}, t_j)$ be the solution computed after time t_j. Here $\mathbf{x}$ is the vector containing the discretization of the domain $[-L, L]$ into N points. To ensure that the gradient flow has stabilized, we compute over M time

spans of fixed length (M is variable), each time continuing the gradient flow, and terminate when one of the following three conditions is satisfied:

(1) Define

$$U(\mathbf{x}, t_j) := \frac{u(\mathbf{x}, t_j)}{\|u(\mathbf{x}, t_j)\|_{L^\infty}},$$

and

$$D = \frac{1}{2L} \sum_{i=1,\dots,N} [U(x_i, t_M) - U(x_i, t_{M-1})]^2.$$

Then terminate when $D < 0.05$.

(2) Set $H := m/(2L)$, the maximum of a completely vanishing function. Then the flow is terminated when $\|u(\mathbf{x}, t_M)\|_{L^\infty} > 3H$. This is particularly useful when we simulate solutions with small mass.

(3) Terminate when $\|u(\mathbf{x}, t_M)\|_{L^\infty} < 1.3H$.

Conditions (1) and (2) refer to non-trivial solutions, (3) refers to a vanishing solution. The third possibility in the Concentration-Compactness Lemma, dichotomy, has never been observed in any computation, and is therefore not dealt with here. Figure 5.3 is now made by colouring gray all grid points for which the gradient flow terminated with one of the conditions (1) and (2), and colouring white those that terminated with condition (3).

Figure 5.3 indeed illustrates that minimizers for Problem (A) only exist for m sufficiently large. Moreover, as indicated by (5.25) in Remark 5.27, for $\alpha \leq \frac{1}{2(1+c_0)}$ minimizers may not exist at all. This is also supported by Figure 5.3.

In the light of the exponentially slow aggregation of two lumps of mass discussed above, initial conditions for the gradient flow are taken with compact connected support. Furthermore, it proves crucial to take initial conditions close enough to a minimizer. Taking random initial data often does not give representative results. For relatively small mass, good candidates for such initial conditions are block functions $H(x + m/2) - H(x - m/2)$ where $H(x)$ denotes the Heaviside function.

A family of unstable stationary points as initial conditions

In [11] a family of semi-explicit solutions of minimization Problem (A) (or, more precisely, a slight variation of Problem (A)) was constructed. It was also proved that these solutions are unstable, i.e., for any stationary point u of this family there exists a perturbation w such

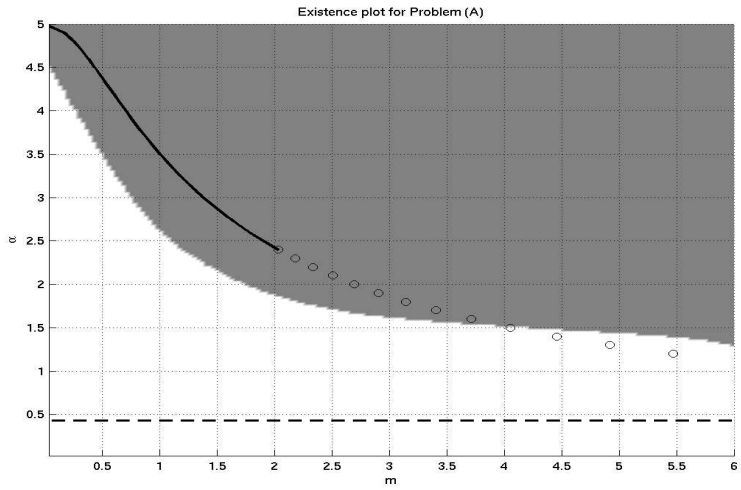


Figure 5.3: Numerical existence of minimizers for Problem (A) using the gradient flow (5.39). The gray region refers to existence, and the white region to non-existence. See the text for more details. A family of unstable stationary points defined by (5.40) is plotted with the solid black line. Where this line continues as circles, the upper contact condition $u + \tau_h u \leq 1$ is violated. The horizontal dashed line corresponds to the bound on α given in Remark 5.27 below which minimizers have been proved not to exist. In all simulations, $c_0 = 0.1, N = 100, h = 10, L = 30$.

that $H''(u) \cdot w \cdot w < 0$. We briefly review the construction of these solutions. Suppose u is a stationary point of Problem (B) with $\operatorname{supp} u = \mathbb{R}$. Then u satisfies

$$\log(u/c_0 + 1) - 2\alpha u \, \kappa * u = \lambda \text{ on } \mathbb{R}.$$

Since $u(x) \to 0$ as $x \to \pm\infty$, we deduce $\lambda = 0$. Setting $\phi = \kappa * u$, we have $-\phi'' + \phi = u$, and hence

$$-\phi'' + \phi = c_0(e^{2\alpha\phi} - 1) \text{ on } \mathbb{R}.$$

Multiplying by ϕ' and integrating over $(-\infty, x)$, we obtain

$$\frac{1}{2}\phi'^2 + G(\phi) = 0, \tag{5.40}$$

where

$$G(\phi) = -\frac{1}{2}\phi^2 + \frac{c_0}{2\alpha}(e^{2\alpha\phi} - 1 - 2\alpha\phi).$$

The constant of integration is determined by requiring that $\phi(x) \to 0$ as $x \to \pm\infty$. If $2\alpha c_0 < 1$ then there exist solutions homoclinic to zero, which have been proved to be unstable (Thm. 8 in [11]).

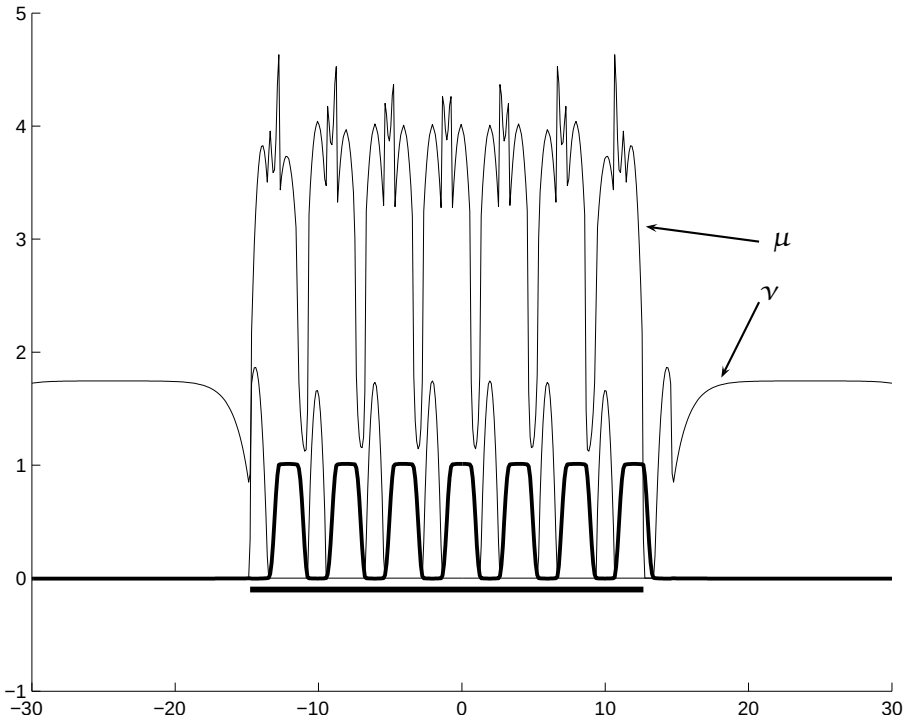


Figure 5.4: An example minimizer of Problem (A), computed with the gradient flow (5.39). The horizontal solid line below the solution signifies the part of the solution where $u + \tau_h u = 1$. The Lagrange multipliers μ and ν are also indicated. Parameters are $\alpha = 6$, $c_0 = 0.1$, $L = 30$, $h = 20$, $N = 600$, $m = 14$.

For a given α, there exist a unique homoclinic solution of (5.40). As can be seen in Figure 5.3, the unstable homoclinic solutions that are admissible functions for Problem (A), i.e., they satisfy the upper contact condition, lie entirely within the existence region of the (m, α)-plane.

Structure of solutions

This section on numerics is completed by illustrating some properties of minimizing solutions of Problem (A). Figure 5.4 shows a minimizer with parameter values α and c_0 such that $2\alpha c_0 > 1$, the region for which existence for all mass $m > 0$ has been proved. Perhaps the most striking feature of this solution is that there is a complete interval

on which $u + \tau_h u = 1$, and that on parts of this interval, $u = 0$. In other words, $\Omega_u \cap \Omega_d$ is non-empty. Note that on that intersection, the Lagrange multipliers μ and ν are both non-zero.

The solution in Figure 5.4 appears again in Figure 5.5, next to a solution with lower α and lower h. We conjecture that there exists an upper bound, independent of h, on u' on Ω_u. Assuming this, it is clear that by decreasing h, u oscillates faster around $\frac{1}{2}$, and does not reach 0 and 1 as extremes anymore. This is shown in the right solution in Figure 5.5. This effect is amplified by lowering α.

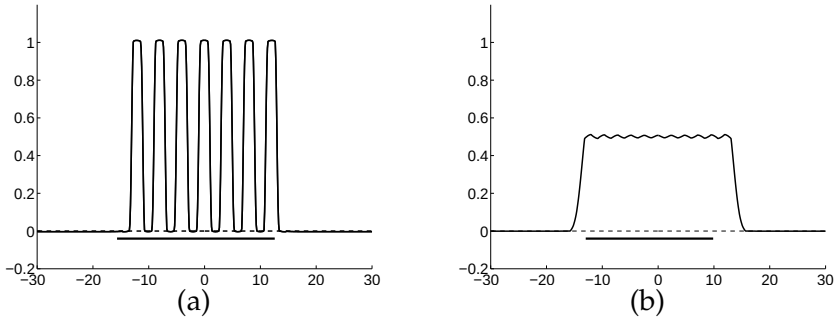


Figure 5.5: Two example minimizers of Problem (A), computed with the gradient flow (5.39). The thin solid line is the solution. The thick solid line signifies the part of the solution where $u + \tau_h u = 1$. Note that in (a) there is a non-empty intersection of the lower and upper contact sets, Ω_d and Ω_u, while in (b) this intersection is empty. In these simulations, the initial condition is a block function with mass m, and height 0.45, positioned in the middle of the domain. Parameters common to both plots are $c_0 = 0.1$, $L = 30$, $N = 600$, $m = 14$. Furthermore, for (a), $\alpha = 6$, $h = 20$, and for (b), $\alpha = 3$, $h = 12$. h is measured in grid points.

For large values of α, minimizers seem to have a non-empty intersection of the lower and upper contact sets, as illustrated in Figure 5.4, and constitute a number of 'humps'. Moreover, numerics suggest that they are symmetric—perhaps a counter-intuitive property given the non-local nature of the upper contact condition.

Since the energy functional favours concentration of mass, we may wonder how these solutions change qualitatively if we increase mass. Figures 5.6 and 5.6 are the result of the following numerical experiment. Let $m_1 < \ldots < m_n$ be given. For each m_i, $i = 1, \ldots, n$, we compute a numerical minimizer with the gradient flow, starting from a block function with height 0.45 centred at 0. Figure 5.6 indicates that one passes

a saddle point, in the passing of which the number of tall humps increases by one. Example solutions before, close to and after the saddle point are illustrated in Figure 5.7. If we keep increasing mass such saddle points are encountered periodically.

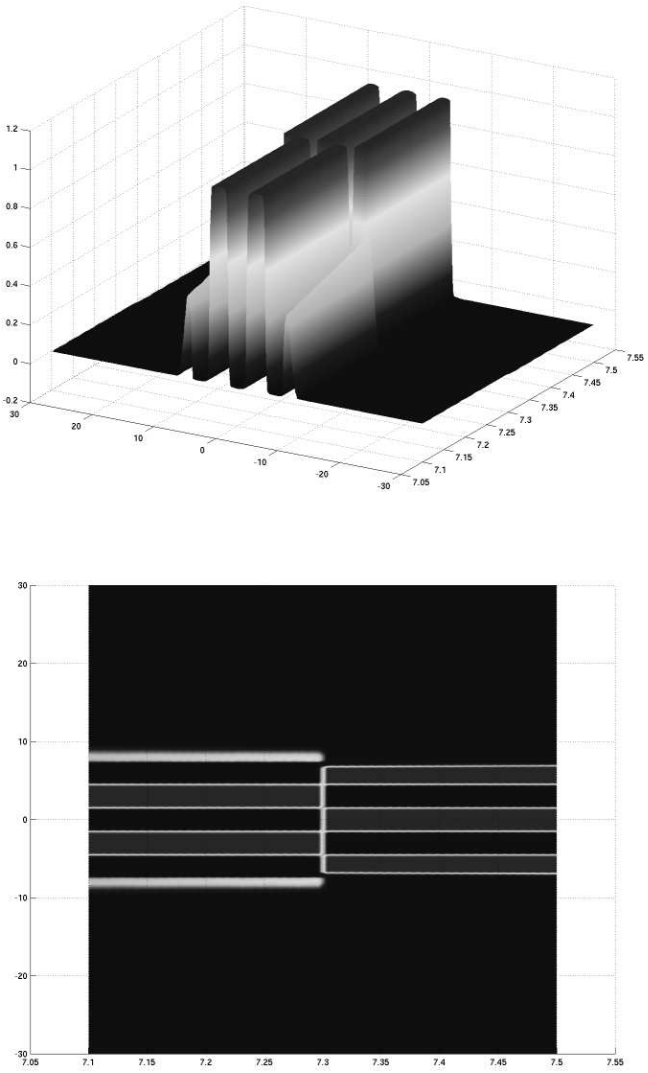


Figure 5.6: While increasing mass but keeping the other parameters fixed one passes a saddle point, manifested in a change in the number of humps of corresponding minimizers (left: 3d view; right: top view). At $m = 7.3$, the two humps that have grown on either side for $m < 7.3$, disappear and a three-hump minimizer is formed. This process occurs for larger mass as well. Single minimizers are illustrated in Figure 5.7. Parameters are as in Figure 5.4, except $h = 30$ and $m \in [7.1, 7.5]$.

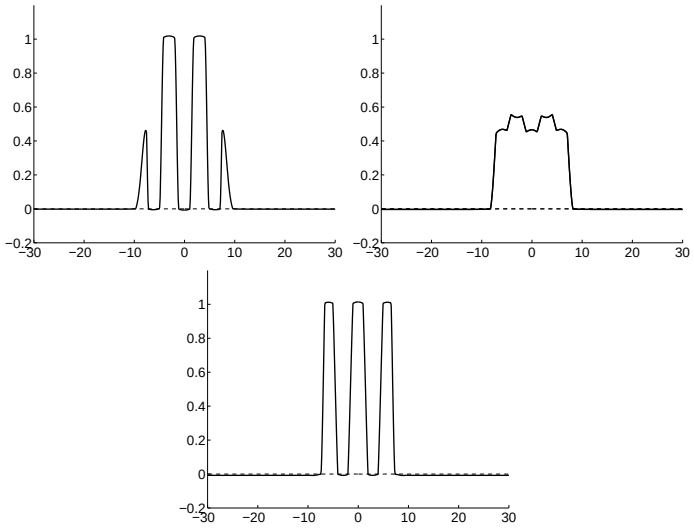


Figure 5.7: Illustration of the passing of a saddle point while increasing mass (see also Figure 5.6). Before passing the saddle point (left top), the minimizer has two tall humps and two flanking smaller humps; very close to the saddle (right top), the minimizer start to switch between two and three tall humps; having passed the saddle (bottom), the minimizer has three tall humps. Parameters are as in Figure 5.4, except $h = 30$, and $m = 7.16$ (left top), 7.32 (right top), and 7.56 (bottom).

5.9. Appendix

THEOREM 5.39. *Suppose for given $\alpha > 0$, $c_0 \in (0, 1/2)$, that there exists an $\bar{m} > 0$ such that for all $m \in (0, \bar{m}]$ there exists a minimizer in K^B_m. Then $\bar{m} = \infty$.*

PROOF. Let $V = (0, \bar{m}]$, and suppose $M \notin V$. Then w.l.o.g. $M < 2\bar{m}$. We use the notion of concentration functions [83], used in the proof of the Concentration-Compactness Lemma [86]. Let $\{u_n\}$ be a minimizing sequence in K^B_M, and define

$$Q_n(t) := \sup_{y \in \mathbb{R}} \int_{y-t}^{y+t} u_n.$$

For any $n \in \mathbb{N}$, $\lim_{t\to\infty} Q_n(t) = M$. Then for any $t > 0$, after possibly extracting a subsequence, the limit for $n \to \infty$ is well-defined (see [86] for more detail). So set

$$Q(t) := \lim_{n\to\infty} Q_n(t).$$

Then

$$\lim_{t\to\infty} Q(t) =: \alpha \in [0, M].$$

If $\alpha = M$, then the sequence was compact, and $M \in V$, a contradiction.

If $\alpha = 0$, then $\{u_n\}$ vanishes. Set $M = \bar{m} + \varepsilon$. Then $\bar{m}, \varepsilon \in V$, and hence there exist functions $u_{\bar{m}} \in K^B_{\bar{m}}$ and $u_\varepsilon \in K^B_\varepsilon$ such that $H(u_{\bar{m}}) = I(\bar{m})$, and $H(u_\varepsilon) = I(\varepsilon)$. Moreover, $u_{\bar{m}}$ and u_ε both have compact support by Corollary 5.23. Hence, after an appropriate translation, u_m and u_ε have disjunct supports. Set $u_M := u_{\bar{m}} + u_\varepsilon \in K^B_M$. Therefore

$$\begin{aligned} H(u_M) - H(u_{\bar{m}}) - H(u_\varepsilon) &= -\alpha \int [u_M \,\kappa{*}u_M - u_{\bar{m}} \,\kappa{*}u_{\bar{m}} - u_\varepsilon \,\kappa{*}u_\varepsilon] \\ &= -\alpha \int [u_{\bar{m}} \,\kappa{*}u_\varepsilon + u_\varepsilon \,\kappa{*}u_{\bar{m}}] < 0. \end{aligned} \tag{5.41}$$

This yields

$$I(M) \le H(u_M) < I(\bar{m}) + I(\varepsilon) < 0.$$

But $\{u_n\}$ is a vanishing sequence, and therefore by Lemma 5.14, $\lim_{n\to\infty} H(u_n) = 0$. This is a contradiction.

So $M \in (0, M)$. Following the proof of the Concentration-Compactness Lemma in [86], we find a sequence $\{y_n\} \subset \mathbb{R}$, such that

$$\bar{u}_n(\cdot) := u_n(y_n + \cdot)$$

is compact in K_α^B. More precisely, for all $\varepsilon \in (0, \alpha/2)$ there exists an $\mathbb{R} < \infty$ such that

$$\int_{-R}^{R} \bar{u}_n(x)dx > \alpha - \varepsilon.$$

By Remark 4.6 we may assume that $\{\bar{u}_n\} \subset L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and hence there exists a $v \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that $\bar{u}_n \rightharpoonup v$. Moreover, $H(\bar{u}_n) \to H(v)$. As in the proof of Theorem II.1 in [86], we need only to verify that

$$\int \bar{u}_n \,\kappa * \bar{u}_n \to \int v \,\kappa * v,$$

to conclude $\int h(\bar{u}_n) \to \int h(v)$, and consequently, $\bar{u}_n \to v$ in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Since $\alpha \in (0, M)$, $\{\bar{u}_n\}$ is dichotomous, and therefore there exist sequences $\{u_n^1\}$ and $\{u_n^2\}$ such that for all $\varepsilon > 0$ there exists $n_0 \geq 1$, satisfying for all $n \geq n_0$,

$$\begin{cases} \left\| u_n - (u_n^1 + u_n^2) \right\|_{L^1} \leq \varepsilon, \\ \left\| u_n^1 - \alpha \right\|_{L^1} \leq \varepsilon, \\ \left\| u_n^2 - (M - \alpha) \right\|_{L^1} \leq \varepsilon, \\ d(\operatorname{supp} u_n^1, \operatorname{supp} u_n^2) \to \infty \text{ as } n \to \infty. \end{cases}$$

By Riesz's Lemma 5.21, we may assume that $\bar{u}_n$ is a symmetrized sequence. Hence either u_n^1 vanishes or u_n^2 vanishes. Without loss of generality, we assume the latter. But then u_n^1 is a compact sequence, and, following Theorem II.1 in [86] again, we can indeed conclude

$$\int u_n^1 \,\kappa * u_n^1 \to \int v \,\kappa * v,$$

and $u_n^1 \to v$. In all, we find a function $v \in K_\alpha^B$ such that $H(v) = I(M)$, and $\alpha \in V$, since v can also be seen as the limit of a compact sequence in K_α^B by restricting u_n^1 to make functions of mass α. Moreover, $I(\alpha) = H(v)$.

Now, if $\alpha \in (0, \bar{m})$, then by Remark 5.31 we have $I(M) < I(\bar{m})$. But by that same remark, $I(M) = I(\alpha) < I(\bar{m})$, a contradiction.

This leaves $\alpha \in [\bar{m}, M)$. But now we again obtain a contradiction: since $M < 2\bar{m}$ and $M = \alpha + \varepsilon'$ we know $\alpha, \varepsilon' \in V$. But then

$$I(M) < I(\alpha) + I(\varepsilon') < I(\alpha).$$

This completes the proof. □

We cannot expect this method of proof to work for a general class of minimization problems. Suppose the minimizing sequence $\{u_n\}$ splits

into two sequences u_n^1 and u_n^2 as in the definition of dichotomy. If we may not use symmetrization, it may well hold that for all $L, M < \infty$,

$$\left(\lim_{n\to\infty} \sup_{y\in\mathbb{R}} \int_{y-L}^{y+L} u_n^1\right) \cdot \left(\lim_{n\to\infty} \sup_{z\in\mathbb{R}} \int_{z-M}^{z+M} u_n^2\right) > 0. \tag{5.42}$$

In that case certainly

$$\int u_n \,\kappa{*}u_n \not\rightharpoonup \int v \,\kappa{*}v,$$

where v is the weak limit of u_n. For now, it remains an open problem to show that (5.42) is excluded for minimizing sequences for Problem (A).

CHAPTER 6

The adaptiveness of defence strategies against cuckoo parasitism

Defence strategies against predators or more specifically parasites are ubiquitous and diverse [52, 56]. Given the many marvellous adaptations known in the natural world we may sometimes wonder why a certain obvious strategy has not evolved. The main question under investigation in this chapter is: why is chick rejection so scarce among cuckoo hosts? One seemingly contradictory situation in defence mechanisms is known from ornithology: although host birds often have great abilities to discriminate against Eurasian Cuckoo *Cuculus canorus* eggs, there seems to be a nearly complete lack of defence when it comes to cuckoo chicks. Why is the defence strategy of chick rejection so scarce among the cuckoo's hosts?

6.1. Modeling

We start by explaining the motivation for our approach. Traditionally, research has focused on egg rejection (see [143] for a short review and [33] for a detailed account). Both experimental approaches and modeling efforts efforts have contributed to the understanding of this phenomenon.

Because of the near complete lack of chick-rejection behaviour by all the cuckoo's host species, it is difficult to set up experiments to test any hypotheses concerning chick rejection. Nevertheless, a number of things have been clarified. It has been shown that some of the cuckoo's

This chapter is based on [109] and [61].

hosts accept chicks of other species too [34, 35], which contradicts earlier suggestions that the cuckoo chick must manipulate its foster parents [38]. However, the cuckoo chick seems to be able to persuade its parents to bring enough food by calling excessively [37].

According to the prevailing view the crucial factor preventing the evolution of chick discrimination has been the absence of the host's own chicks for comparison [79, 34, 87, 111, 33]. The idea that discrimination is easier if there is a model for comparison is supported by the finding that host nestling discrimination has until now only been observed in non-evicting parasites [104, 46, 85]. In the three host-parasite systems in question, hosts can compare their own nestlings with alien ones [34].

However, to discriminate effectively, birds do not always need comparative material: hosts of parasitic birds are able to reject the entire parasite clutch (exchanged by experimenters for an original one) even without any of their own eggs present [89, 138, 118, 80]. More importantly, estrildids (family *Estrildidae*) can discriminate against a whole brood of another species with no conspecific young for comparison [107]. Furthermore, there are non-comparative cognitive systems that could well work in the context of parasitic chick discrimination—discrimination can be innate (mate recognition [125]) or based on the individual's own phenotype (self-referent phenotype matching [63]). A future host can also learn when it was a chick from the begging sounds and an appearance of its own nestmates.

We will now develop a model to gain insight into the evolutionary aspects of chick rejection. Lotem [87] assumed chick rejection to be a trait learned by imprinting upon the chicks in the nest, and therefore concluded that there is a simple explanation for the observed lack of chick rejection: this type of defence strategy is never selectively advantageous if the probability of making a discrimination error is non-zero. However, if we assume that defensive behaviour against cuckoo chicks is an innate trait, we are still at a loss [121]. It is this side of the problem we address in this chapter.

To motivate the choice of our model we briefly look into previous models for egg rejection. Early models have focused either on population dynamics [97] or on population genetics [117, 76, 16]. Takasu *et al.* [133] and Takasu [132] have been the first to include both ecological and evolutionary aspects of the problem. This more dynamic approach was seen to be necessary following the observation of rapid changes in the defence behaviour by one of the cuckoo's hosts in Japan

[100]. Their model has been effective in giving plausible and experimentally testable explanations of two phenomena. First, it has shown that changes in egg-rejection behaviour may be caused by a change in parasitic pressure [133]. Second, it has clarified the difference in distribution of egg-rejection behaviour among hosts of the Eurasian Cuckoo versus those of the Brown-headed Cowbird, *Molothrus ater*, on the basis of their specialist vs. generalist parasitic traits [132]. In our attempt to explain the lack of chick rejection we have a slightly more complicated situation than in the case of egg rejection: the latter already exists among many of the cuckoo's hosts. Hence we have to investigate whether chick rejection could invade a population that already exhibits defensive behaviour towards cuckoo eggs. Is the cuckoo-host system exhibiting an evolutionary equilibrium or an evolutionary lag? It is hence necessary to include at least three host types: hosts that accept both eggs and chicks, hosts that reject eggs and hosts that reject chicks. For completeness, we shall also include a host type that rejects both eggs and chicks.

Before we introduce our model we will make some general remarks. As we have seen the cuckoo has many hosts but each gens generally specializes on only one or perhaps two. Therefore, in this paper we focus on one gens and its one host. Contrary to for instance Lotem [87] we assume that the host defence systems are determined by hereditary factors. Recent studies on defensive behaviour by cuckoo hosts support this assumption [120]. We treat the various host types as separate species of birds. We combine a clonal model for the four host types with two predator-prey equations for the interaction between cuckoos and hosts to investigate the brood parasite system. It is widely assumed that there are costs associated with displaying egg rejection, due to discrimination errors [119, 97, 34, 113, 115, 98, 94]. These egg-rejection costs are assumed to be small but are taken into account in our analysis, and will play an essential role. In this paper we assume that chick-rejection behaviour also entails similar recognition costs. We do not take into account any hereditary variation within the cuckoo population.

Let P_t be the population density of female cuckoos and H_t that of the female hosts in year t. We assume that surviving offspring breed in the year after they are hatched. If the female cuckoo finds a nest she will lay a single egg and the chick will grow up with a constant probability G to survive to the next breeding season. The adult female cuckoo survives to the next season with a probability s_P. Similarly, we

introduce a constant s_H which measures the intrinsic survival rate of adult host birds. Here, as in the rest of the paper, we neglect any effects due to intra-specific competition.

We distinguish four host types: all-accepters, egg-rejecters, chick-rejecters and all-rejecters. Their frequencies in the total host population are denoted by h_t^a, h_t^e, h_t^c, h_t^{ec} respectively, which add up to one.

The cuckoo is assumed to perform a random search with a search efficiency measured by a parameter a, called the area of discovery by Nicholson and Bailey [103]. The probability that a host nest escapes from parasitism is thus given by e^{-aP_t}, the zeroth term in a Poisson distribution [97]. The density of cuckoos in the next generation is

$$P_{t+1} = s_P P_t + (1 - e^{-aP_t}) H_t G(h_t^a + q_e h_t^e + q_c h_t^c + q_{ec} h_t^{ec}), \tag{6.1}$$

see also [133]. The first term corresponds to the surviving adults, the second to the successfully raised young from the nests that have not escaped parasitism. We will introduce the constants q_e, q_e, and q_{ec} in a moment, and explain their occurrence in equation (6.1) at the end of the section.

The total density of offspring in the host population is a sum over the contributions from the different host types:

$$H_t\left(f_a h_t^a + f_e h_t^e + f_c h_t^c + f_{ec} h_t^{ec}\right).$$

Here we have introduced fitness functions for each of the host types, denoted by f_a, f_e, f_c, and f_{ec}. They will be discussed shortly. In the absence of cuckoo parasitism the number of individuals in a certain area is limited by the available resources. Taking this factor into account with a parameter k, the host density for the next year is given by

$$H_{t+1} = \frac{H_t}{1 + \frac{H_t}{k}}\left(s_H + f_a h_t^a + f_e h_t^e + f_c h_t^c + f_{ec} h_t^{ec}\right). \tag{6.2}$$

We proceed with the fitness functions of the host types. In general all host types will suffer to some extent from an increase in parasitism by the cuckoos, in the sense that they will lose some offspring. So for all host types we assume that the fitness functions are monotonically decreasing functions of P_t. In the absence of parasitism however, we expect some differences in the number of offspring produced by the various host types. Since we have assumed that the rejecting host types make some errors in their attempts to discriminate cuckoo eggs or chicks, we assume that the all-accepting pairs have a slight advantage when the parasitic pressure is low.

Let f be the number of offspring per annum raised by an all-accepting host pair that is not parasitized. If we multiply this by the probability for such a pair to escape parasitism we find the number of offspring for all-accepter pairs in terms of cuckoo density:

$$f_a = fe^{-aP_t}.$$

The corresponding fitness functions for the rejecter types are given by

$$f_e = e_1 fe^{-aP_t} + e_2 f(1 - e^{-aP_t}),$$
$$f_c = c_1 fe^{-aP_t} + c_2 f(1 - e^{-aP_t}),$$
$$f_k = k_1 fe^{-aP_t} + k_2 f(1 - e^{-aP_t}),$$

where $e_1 f$ and $e_2 f$ are the expected number of offspring per annum raised by unparasitized and parasitized egg-rejecters respectively, and similarly for chick-rejecters and all-rejecters.

We shall neglect any physiological costs associated with the behavioural capability for rejection, but shall take account of the costs of recognition errors of types I and II. By type I errors we mean mistakenly ejecting one's own egg or chick in the absence of cuckoo parasitism. Let p_e be the probability that an egg rejecter makes a type I error and removes one of its own eggs by mistake, and b_e the relative pay-off for raising a clutch with one egg removed. Then

$$e_1 = (1 - p_e) + p_e b_e.$$

Similarly,

$$c_1 = (1 - p_c) + p_c b_c,$$

and

$$k_1 = (1 - p_e)(1 - p_c) + b_e p_e (1 - p_c) + b_c p_c (1 - p_e) + b_{ec} p_e p_c,$$

with the obvious notation.

By type II errors we mean that the hosts sometimes fail to spot the cuckoo's eggs or chicks in their nests. Let q_e be the probability of an egg-rejecter making a type II error, i.e., mistakenly accepting a cuckoo egg. The relative pay-off for a host that accepts a cuckoo egg is zero. Hence we have

$$e_2 = (1 - q_e) b_e.$$

We have assumed here that an egg-rejecter has the same pay-off b_e for rejecting one of its own eggs when unparasitized as it does for rejecting a cuckoo egg when parasitized: it merely has one egg less, in the first case removed by itself and in the second by the laying cuckoo. For chick-rejecters this is not the case: when an unparasitized chick-rejecter makes a type I error and ejects one of its own chicks, it still has the rest

of its clutch, and its pay-off is b_c. When it is parasitized and ejects the cuckoo chick, however, its pay-off varies between b_e (loss of the egg removed by the laying cuckoo) and zero (loss of the clutch), depending on the damage done by the cuckoo chick before it is discovered and ejected. We shall set the pay-off to γb_e, where γ is a measure of how much of the clutch is saved on average. Thus

$$c_2 = (1 - q_c)\gamma b_e.$$

For all-rejecters we find

$$k_2 = (1 - q_e)b_e + q_e(1 - q_c)\gamma b_e.$$

Now note that $0 < b_{ec} < b_c, b_e < 1$; it is better to lose either an egg or a chick than to lose both, but it is better still to lose neither. Moreover $b_c \leq b_e$; if one potential offspring is to be lost it might as well be lost early (at the egg stage), so that it no longer requires resources (however minimal). Note also that $0 \leq p_e, p_c, q_e, q_c < 1$, with e.g., $p_e = 0$ if and only if the egg-rejecter never makes a type I error. It follows immediately that

$$0 < e_2 < e_1 \leq 1, \quad 0 < c_2 < c_1 \leq 1, \quad 0 < k_2 < k_1 \leq 1;$$

(i) even rejecters are disadvantaged by parasitism, because the cuckoo ejects one of their eggs when laying its own, and (ii) rejecters are no better off than accepters in the absence of parasitism (and are worse off unless they never make a type I error). It also follows immediately that

$$k_1 \leq e_1, c_1, \quad k_2 > e_2, c_2;$$

(iii) rejecters with both defence strategies are no better off than those with only one in the absence of parasitism (and are worse off unless they never make a type I error), but (iv) they are better off than those with only one if they are parasitized. However, we never see hosts that employ both defence strategies. The explanation of this depends on a subtler argument that we give later. All these inequalities, and hence conclusions (i) to (iv), hold for any allowable parameter values. The relationships between e_1, c_1, e_2 and c_2, on the other hand, depend on the particular parameter values chosen, and in particular on the value of γ. Unless the biologically unreasonable assumption is made that the cuckoo chick is always discovered before it does any damage then $\gamma < 1$. This will play an important part in our explanation of why egg-rejecters rather than chick-rejecters are observed in nature.

An illustration of a typical set of fitness functions can be found in Figure 6.1.

With these fitness functions we may determine the host type frequencies in the next generation:

$$h^a_{t+1} = h^a_t \frac{s_H + f_a}{D}, \tag{6.3}$$

$$h^e_{t+1} = h^e_t \frac{s_H + f_e}{D}, \tag{6.4}$$

$$h^c_{t+1} = h^c_t \frac{s_H + f_c}{D}, \tag{6.5}$$

$$h^{ec}_{t+1} = h^{ec}_t \frac{s_H + f_{ec}}{D}, \tag{6.6}$$

where $D = s_H + f_a h^a_t + f_e h^e_t + f_c h^c_t + f_{ec} h^{ec}_t$.

We may now explain the cuckoo equation (6.1) completely: the defensive host types contribute to the next generation of cuckoos if they have failed to discriminate the cuckoo's eggs or chicks. This amounts to the factors q_e, q_c and q_{ec} found in the equation.

As a remark, note that a can be scaled out of the equations, but this is not done for two reasons: it reduces the number of parameters only by one giving only a small gain, and the current parameter has a well-defined biological interpretation, contrary to its rescaled counterpart.

Equations (6.1), (6.2), and (6.3) to (6.6) constitute our model.

6.2. Analysis

As an introduction to the characteristics of the model we will present a simple, intuitive analysis of a general cuckoo–host system. We believe that this intuitive analysis is helpful, at this stage, even though the full model exhibits behaviour that is both different and more complex than that suggested by the simple intuitive argument. Moreover, the simple argument does include the most important clue to answering the main question discussed in this paper. We will therefore treat it as a background against which we present the more detailed analysis below.

> At low cuckoo densities the all-accepting hosts have highest fitness, whereas the defending host types are fitter for high cuckoo densities. Hence, if the cuckoo density were to remain low (or high), the frequency of the accepting (defending) host type will approach unity. Let us assume that the system will converge to an equilibrium solution, and assume the following two dynamical properties of the model:

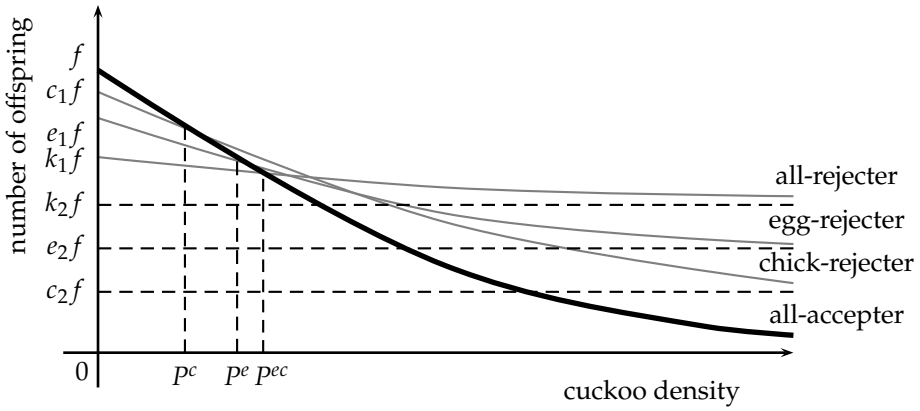


Figure 6.1: The average number of female offspring for each of the host types, all-accepter, egg-rejecter, chick-rejecter and all-rejecter pairs, respectively. The three values P^e, P^c and P^{ec} are the cuckoo densities at which the fitness of the all-accepter hosts equals the fitness of egg-, chick- and all-rejecter hosts respectively. In the diagram they satisfy $P^c < P^e < P^{ec}$, but this order depends on the parameters of the problem. When P_t is small (less than P^c) the accepter pairs produce more offspring, but for higher cuckoo densities the defending hosts are reproductively fitter than accepters. Note that the discrimination costs for all-rejecters is proportionally higher than for either egg or chick rejection, but that these hosts are fitter under high parasitic pressure. The dashed horizontal lines are the asymptotic values for each of the defensive fitness functions for large P_t.

> 1. If there are only accepting hosts and the cuckoo numbers are low, the cuckoo numbers will increase.
> 2. If there are only defending hosts and the cuckoo numbers are high, the cuckoo numbers will decrease.
>
> Then we expect that there is an equilibrium value for the cuckoo density at which the fitnesses of defending and accepting hosts are equal.

Numerical investigation shows that this argument gives a rough description of the characteristics of the model. We give an illustration of the effect described in the argument in Figure 6.2. However, this argument misses additional behaviour: we may encounter for instance quasi-periodic solutions, or stable coexistence of two defensive

host type. We start with a discussion of the solutions described by the heuristic argument.

Intermediate rejection frequencies

In certain parameter value ranges we find one of three equilibrium solutions which we denote by S^e, S^c and S^{ec}. At S^e for instance, we find a steady state with coexistence of all-accepter hosts and egg-rejecters, i.e., with $h^a + h^e = 1$ and both frequencies positive. Analogous descriptions can be given for the other solutions S^c and S^{ec}. The explicit analytic derivation of the solution is given in the Appendix.

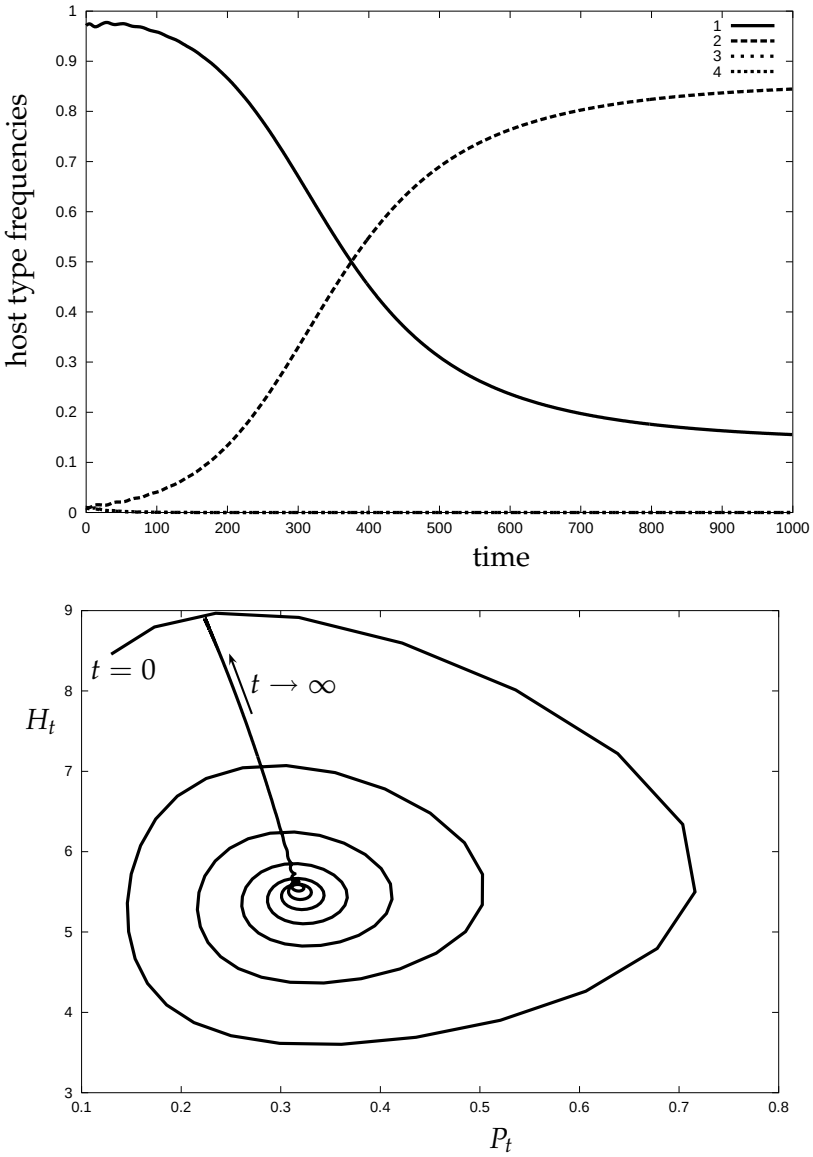


Figure 6.2: A typical numerical simulation of the model, illustrating damped oscillations converging to a steady state in which egg-rejecting (curve 2) and all-accepting hosts (curve 1) coexist. This equilibrium is called S^e. Note that chick- rejecters (curve 3) and all-rejecters (curve 4) increase very briefly but are 'outcompeted' by the fitter egg-rejecting hosts and drop to zero almost immediately. The bottom picture shows the temporal dynamics of cuckoos (curve 5) with respect to hosts (curve 6). Parameters used are $s_H = 0.5$, $s_P = 0.5$, $f = 0.7$, $a = 0.7$, $G = 0.15$, $k = 90$,

$e_1 = 0.95$, $e_2 = 0.3$, $q_e = .5$, $c_1 = 0.9$, $c_2 = 0.15$, $q_c = .5$, $k_1 = 0.85$, $k_2 = 0.375$, $q_{ec} = .75$ and initial conditions $P_0 = 0.1$, $H_0 = 8$, $h_0^a = 0.97$, $h_0^e = h_0^c = h_0^{ec} = 0.01$. In all numerical investigations, we have only changed s_H, f, k, and the discrimination costs. The parameter values conform to those of Takasu *et al.* (1993).

Here we merely state that they are of the form

$$S^e = (P^e, H^e, h^a, h^e, 0, 0), \tag{6.7}$$
$$S^c = (P^c, H^c, h^a, 0, h^c, 0), \tag{6.8}$$
$$S^{ec} = (P^{ec}, H^{ec}, h^a, 0, 0, h^{ec}), \tag{6.9}$$

where P^e denotes the steady state value of P_t in the case of intermediate egg rejection, etcetera. The equilibrium values for P_t can be determined directly from the fitness functions: at each of these solutions the fitness of all-accepters is equal to the fitness of the respective defensive host type. So in the case of egg-rejecters, we solve $f_a = f_e$ for P_t to find the desired result.

This result raises the question, to which of these solutions will the system converge? Numerical investigations have shown that there is a simple rule to determine this: one computes the equilibrium cuckoo densities P^e, P^c and P^{ec} and determines which is the smallest. The equilibrium solution corresponding to this cuckoo density then is the one to which the system will converge. This is a direct application of the heuristic argument stated at the beginning of this section, given that the fitness functions are monotonically decreasing functions in P_t: for P_t less than the smallest equilibrium value, P_t will increase; for P_t larger than the smallest equilibrium value, the all-accepters will have highest fitness and will cause a decrease of P_t.

We have assumed a number of things in the preceding discussion, any of which may under certain conditions be violated and give additional behaviour not explained by the argument above:

- The system will actually converge to an equilibrium solution.
- In such instances S^e, S^c, or S^{ec} are the only possible steady states.
- Among P^e, P^c, and P^{ec} there is one and only one value which is strictly smaller than the other two.

Each of these assumptions does not have to hold, giving qualitatively different behaviour. These phenomena will now be explored in more detail.

Extending the three equilibria

The three equilibria S^e, S^c and S^{ec} may be extended in a natural way to include limit cases. We treat these extensions as separate equilibria since these extensions are not described by the heuristic argument, they are treated as separate equilibria. We describe the occurrence of the following cases: absence of cuckoos, a cuckoo population in coexistence with a completely all-accepting host population, and a cuckoo population in coexistence with a host population that consists of a single defensive host type.

We start by looking at the system in the case of all-acceptance of the host pairs. In general the system will converge to a unique equilibrium solution. The equilibrium solution for the host population in absence of the cuckoo may be found by setting $H_{t+1} = H_t$ and $P_t = 0$ in (6.2). Let us call this steady state $S^0 = (P^0, H^0, h^a, h^e, h^c, h^{ec}) = (0, H^0, 1, 0, 0, 0)$. We then find

$$H^0 = k(s_H + f - 1).$$

This number corresponds to the carrying capacity of the host population. If $H^0 < 0$ the host population goes extinct since the death rate then exceeds the birth rate ($1 - f > s_H$). From now we will assume that $1 - f < s_H$.

For larger values of k we find that in the absence of any rejecter hosts the system evolves towards an equilibrium solution $S^1 = (P^1, H^1, 1, 0, 0, 0)$, where P^1 and H^1 are the unique solution of

$$\begin{cases} H^1 &= \frac{(1-s_P)P^1}{G(1-e^{-aP^1})}, \\ H^1 &= k(fe^{-aP^1} + s_H - 1). \end{cases}$$

If P^1 approaches zero we find

$$\frac{s_P - 1}{Ga} + k(s_H + f - 1) = 0.$$

Therefore the a critical value of our chosen bifurcation parameter k for the survival of the cuckoo population is given by

$$k^P := \frac{1 - s_P}{aG(s_H + f - 1)}.$$

For smaller values of k the cuckoo population goes extinct. This argument is identical to the analysis in Takasu *et al.* [133].

We now consider the relationship between these two new steady states and S^e, S^c, and S^{ec}. Since $h^e \in (0, 1)$ we may solve $h^e > 0$ from the analytic expression in (6.18) in terms of k to find a minimal value

for k for S^e to be meaningful. Let us call this value k_0^e. Similarly we find minimal values k_0^c and k_0^{ec}. The methods to derive these expressions are given in the Appendix. We note here that $k^P < k_0^e, k_0^c, k_0^{ec}$, thus excluding the possibility that any defensive host types could establish themselves before the cuckoos were present. When k is in the interval $(k^P, \min\{k_0^e, k_0^c, k_0^{ec}\})$ the system converges to S^1.

Intuitively the occurrence of these values k_0^e, k_0^c and k_0^{ec} can be explained by our expectancy that the cuckoo has to search well enough, and the environmental carrying capacity for the hosts has to be sufficiently high.

We can, on the other hand, also find maximal values for k for these equilibria S^e, S^c and S^{ec} to be converged upon. This can be done by solving for instance $h^e < 1$ in the case of S^e. We refer to the Appendix for the formal calculations. We denote these values by k_1^e, k_1^c and k_1^{ec}. We can now formulate more precisely when one of the three equilibria may be converged upon, which we illustrate again in the case of S^e: S^e may be attained if $k \in (k_0^e, k_1^e)$ and P^e is the smallest of the three equilibrium values P^e, P^c and P^{ec}. Completely analogous conditions can be given for S^c and S^{ec}.

We thus find natural extensions for all $k > 0$ of the equilibria S^e, S^c and S^{ec}. As k increases we go through four stages:

- Only hosts and no cuckoos, $k \in (0, k^P)$.
- Coexistence of cuckoos with all-accepting hosts, $k \in (k^P, k^*)$, where $k^* = \min\{k_0^e, k_0^c, k_0^{ec}\}$.
- Coexistence of cuckoos with a stable mixed population of all-accepting and one defensive host type, $k \in (k^*, \hat{k})$, where $\hat{k}$ is the corresponding k_1^e, k_1^c or k_1^{ec}.
- Coexistence of cuckoos with a host population existing only of hosts of one defensive type, $k \in (\hat{k}, \infty)$, where $\hat{k}$ is one of $\{k_1^e, k_1^c, k_1^{ec}\}$.

This is illustrated in Figure 6.3 in the case of egg-rejecters.

Here we have still assumed two things: the system converges to an equilibrium and the steady state in which one defensive type is the only host type is stable for all $k > \hat{k}$ where $\hat{k}$ is one of $\{k_1^e, k_1^c, k_1^{ec}\}$. The violation of the latter condition is discussed later. When the first assumption is not valid we can find quasi-periodic solutions. These are discussed in the next section.

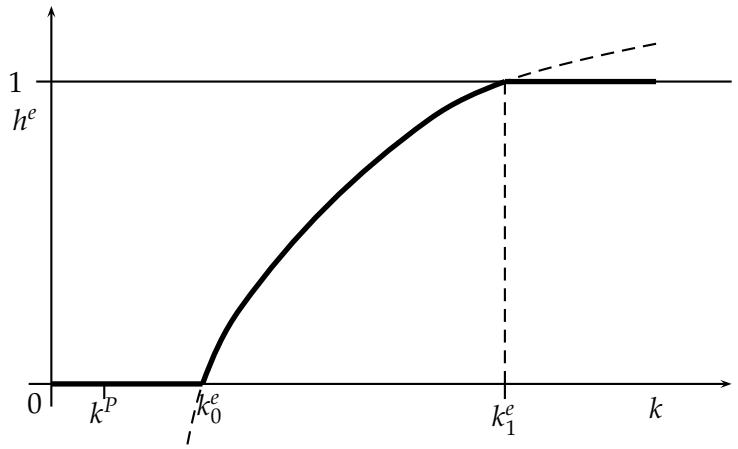


Figure 6.3: The coexistence of egg-rejecters and all-accepters at steady state S^e (bold curved line, and dashed its hypothetical extension below $h^e = 0$ and above $h^e = 1$). The extension of S^e to all $k > 0$ is indicated by the horizontal bold lines. Starting at $k = 0$ we find no cuckoos present until k^P, cuckoos but no defensive hosts between k^P and k_0^e, cuckoos and a mixed defensive and accepting host population between k_0^e and k_1^e and cuckoos and only egg-rejecters for $k > k_1^e$. This figure corresponds to curve (1) of Figure 6.4.

No convergence to an equilibrium solution

Even when the conditions for any of the three steady states is satisfied, the system does not have to converge to any of the three. We again illustrate this in the case of S^e. There is an asymptotic value for s_H, s_H^e say, with the property that for $s_H < s_H^e$ the steady state S^e is not attained. The analytical derivation of s_H^e is given in the Appendix. This is illustrated in Figure 6.4. Numerical investigations show that there is an s_H-interval in which S^e is not attained, but the system converges to a quasi-periodic orbit (see Figure 6.5 for an example). These stable oscillations are both in cuckoo and host numbers, and also in host type frequencies. Mathematically their appearance corresponds to a Naimark-Sacker bifurcation, the equivalent of the Hopf bifurcation for discrete time systems (see e.g., [8, p. 261]). In Figure 6.4 the occurrence of this bifurcation has been placed in a broader context of parameters.

The occurrence of this phenomenon seems to be independent of the stability or occurrence of the three equilibria S^e, S^c and S^{ec}. The s_H-interval in which these periodic solutions are found is concentrated around the asymptote $s_H = s_H^e$. Note that we also find non-trivial defensive hosts (i.e., $h_t^e > 0$) for $s_H < s_H^e$, which do not correspond to the intermediate frequency h^e in steady state S^e. This is not predicted by the heuristic argument.

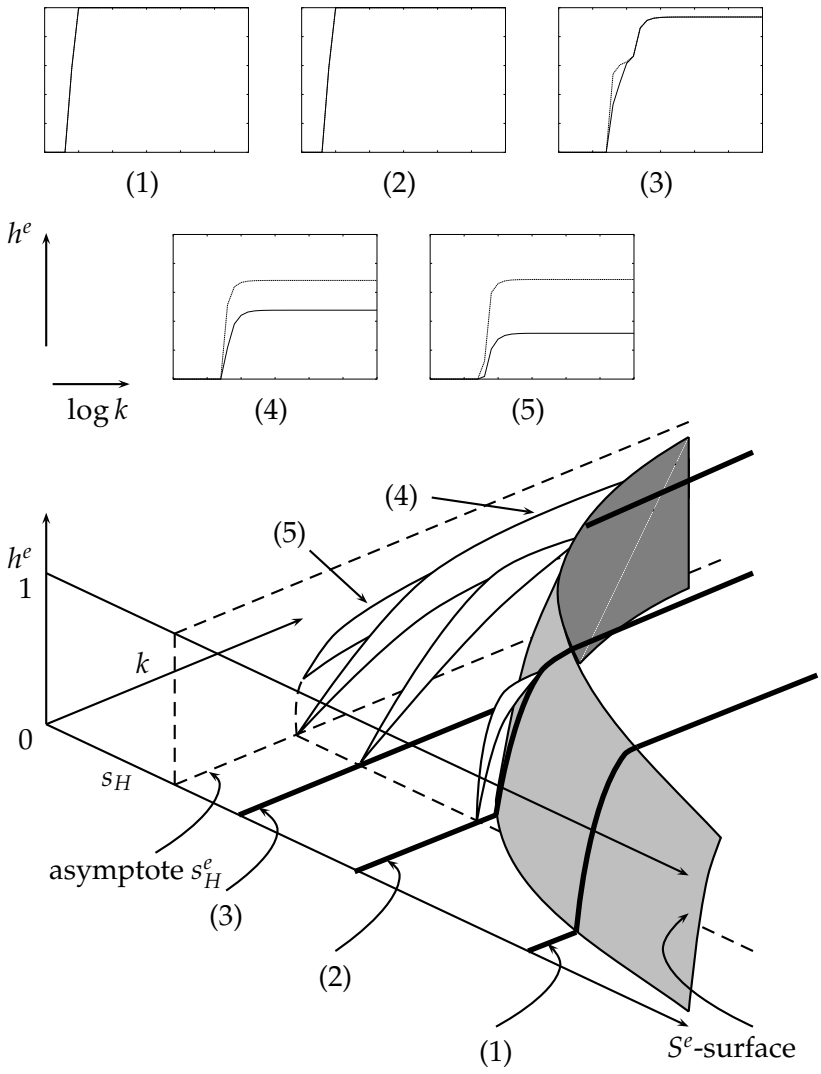


Figure 6.4: Assuming P^e to be the smallest of $\{P^e, P^c, P^{ec}\}$, we find qualitatively different solutions if we vary parameters k and s_H. With s_H large, we find the standard solution $\bar{S}^e$, such as (1), (2) and (3). The region with intermediate frequencies, denoted by S^e forms a surface in the (k, s_H, h^e)-space, which asymptotically converges to $k = 0$ and $s_H^e = 1 - fe_2/(1 - e_1 + e_2)$ in the two planes $h^e = 0$ and $h^e = 1$. For small values of s_H we find quasi-periodic solutions which are confined to the white

curved regions, originating at the dashed line. One implication is that solutions exist for s_H less than the asymptotic value in which there is sustained egg rejection, for instance (5). The width of the oscillating regions is largest in the (s_H, h^e)-plane at the s_H-asymptote, see (4). Keep in mind that for large values of k, $\bar{S}^e$ is only locally stable. Keeping initial conditions fixed, the system can make a transition from egg rejection as steady state to all-rejection, as explained in Figure 6.6. This effect is not shown in the above picture. The five small pictures correspond to the solutions shown in the large picture. Note that in all but the left figure, we find regions where Naimark-Sacker bifurcations have occurred. As in the large picture, Figure 1 does not include the instability of the $h^e = 1$ solution for larger k. Figures 2 and 3 do not show the unstable part of the egg-rejecter equilibria, namely the part between $h^e = 0$ and where it attaches to the closed oscillatory regions. These figures have been made with parameter values as in Figure 6.2 but with, from left to right, $s_H = 0.5$, 0.379, 0.3785, 0.377 and 0.37; $\log k \in (3.8, 8.6)$ (Figure 1) and $\log k \in (3.5, 18.5)$ (Figures 2 to 5).

More coexisting host types

The heuristic argument predicts that the system converges to a steady state in which the cuckoo population is equal to the smallest value of P^e, P^c and P^{ec}. However, for some parameter values we may find a stable coexistence between two defensive host types rather than between one accepting and one defensive host type. As an example, in Figure 6.6 we see that although for smaller values of k we find the familiar coexistence between all-accepters and egg-rejecters, for larger values of k the egg-rejecters and all-rejecters are in coexistence. This steady state of intermediate egg- and all-rejecters will be denoted by $S^{e,ec}$ and is of the form

$$S^{e,ec} = (P^{e,ec}, H^{e,ec}, 0, h^e, 0, h^{ec}).$$

Similarly, we may encounter a steady state $S^{c,ec}$ in which chick-rejecters and all-rejecters are coexisting in intermediate frequencies of the form

$$S^{c,ec} = (P^{c,ec}, H^{c,ec}, 0, 0, h^c, h^{ec}),$$

or an equilibrium $S^{e,c}$,

$$S^{e,c} = (P^{e,c}, H^{e,c}, 0, h^e, h^c, 0),$$

in which we find egg-rejecters and chick-rejecters The analytical derivations of these equilibria are given in the Appendix.

Still other combinations of host types in equilibrium are possible, but we have only found them when two or more equilibrium cuckoo densities are chosen to be equal, e.g., $P^e = P^c < P^{ec}$. In such cases

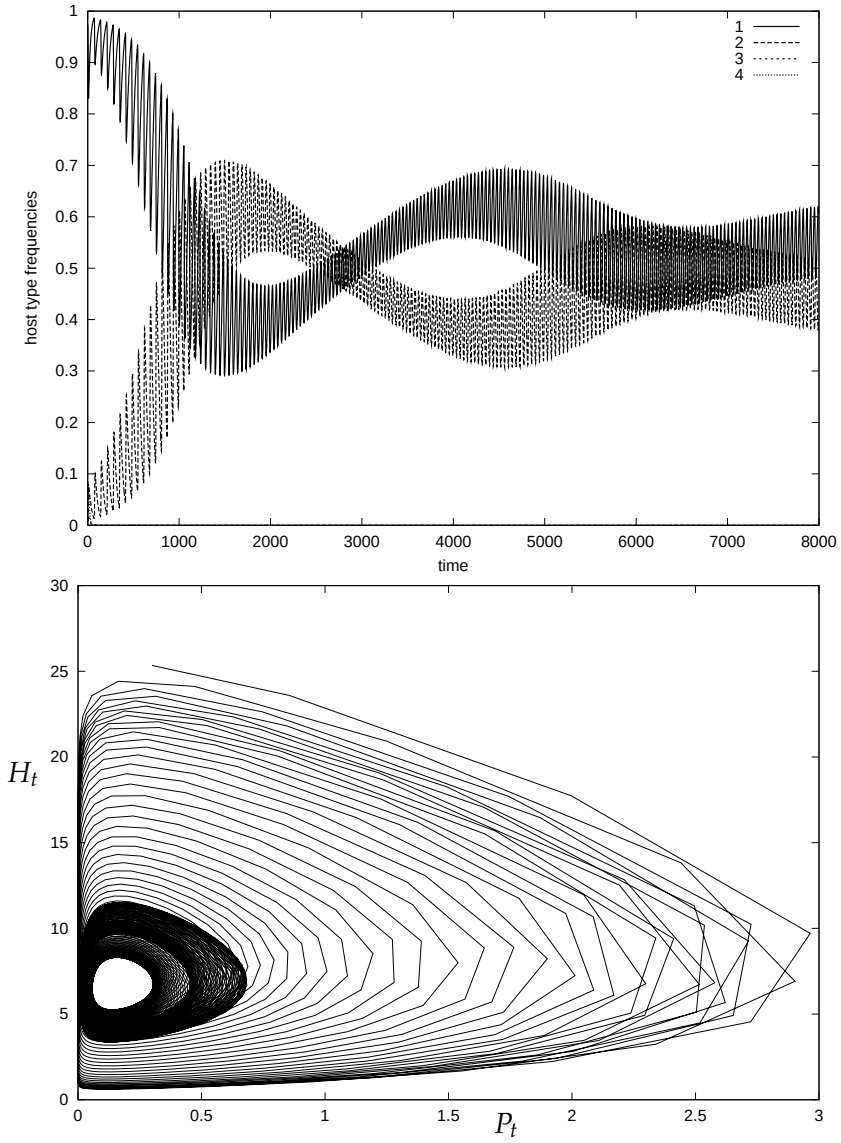


Figure 6.5: Numerical simulation showing quasi-periodic behaviour of both all-accepter and egg-rejecter host types (curves 1 and 2 respectively). The temporal dynamics of the host frequencies are illustrated in the top picture; the cuckoo and host densities are compared in the lower one. Parameter values and initial conditions are identical to Figure 6.2, with the exception of $k = 3000$ and $s_H = 0.378$. Curves 2 and 4 (chick rejection and all-rejection) quickly drop to zero and remain there.

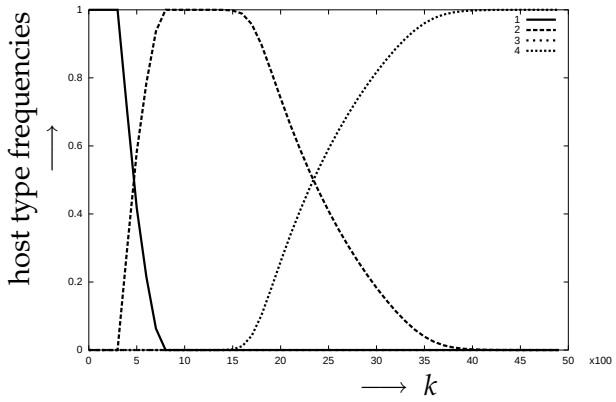


Figure 6.6: For small values of k we find coexistence between all-accepters (curve 1) and egg-rejecters (curve 2), but for larger values the all-rejecters (curve 4) start to increase: we find a steady state $S^{e,ec}$ in which all-rejecters can be found alongside the egg-rejecters. For even larger k the all-rejecters are the only hosts. In all these cases we have $P^e < P^c, P^{ec}$. The chick-rejecters (curve 3) are absent for all k. Parameters used are as in Figure 6.2, but with $f = 0.71212$, $e_1 = 0.882$ and $c_1 = 0.934$.

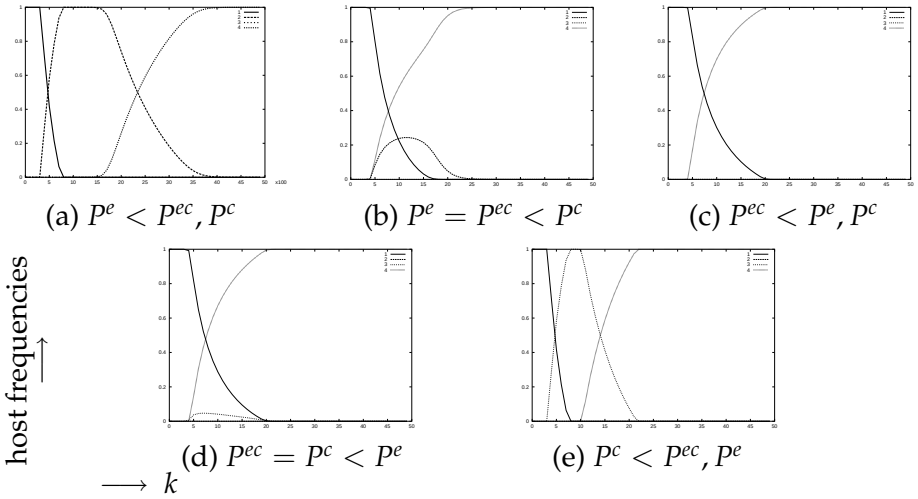
one may find one-parameter families of steady states when two cuckoo densities are the same, or a two-parameter family if $P^e = P^c = P^{ec}$. Since this is biologically implausible, these mathematical properties of the model don't provide additional insight in this biological context.

Multiple fittest strategies

Recall that P^e, P^c and P^{ec} are the three values of P_t at which the fitness of the all-accepters coincides with the egg-rejecters, chick-rejecters and all-rejecters respectively. As we have seen, the heuristic argument may only be applied if there is a unique smallest value among P^e, P^c and P^{ec}. We will now discuss what happens if there are two or more smallest values. We have one of the following four options:

$$P^e = P^{ec} < P^c, \tag{6.10}$$
$$P^c = P^{ec} < P^e, \tag{6.11}$$
$$P^e = P^c < P^{ec}, \tag{6.12}$$
$$P^e = P^c = P^{ec}. \tag{6.13}$$

To understand the behaviour of the model in any of these four cases we investigate the effect of passing through them in parameter space. For instance, in Figure 6.7 we have varied parameters such that we



(a) $P^e < P^{ec}, P^c$ (b) $P^e = P^{ec} < P^c$ (c) $P^{ec} < P^e, P^c$

(d) $P^{ec} = P^c < P^e$ (e) $P^c < P^{ec}, P^e$

Figure 6.7: We make a transition from a system with $P^e < P^{ec}, P^c$ (fig. *a*) to $P^c < P^{ec}, P^e$ (fig. *e*). In Figure *a* we find the coexistence between all-accepters (curve 1) and egg-rejecters (curve 2) for small values of k and between egg-rejecors and all-rejecters (curve 4) for larger values of k. Figure *b* shows the coexistence of three host types, namely all-accepters, egg- and all-rejecters. This is one realization of a one-parameter family of equilibria, and it is dependent on initial conditions. In Figure *c* we have the familiar steady states in which all-accepters and all-rejecters coexist in intermediate frequencies. In Figure *d* we find all-accepters, chick- and all-rejecters incoexistence, and in Figure *e* we have the same behaviour as in Figure *a*, with egg-rejecters replaced by chick-rejecters. Note that all these transitions are non-continuous. Parameters are as in Figure 6.2, but with the exception of $f = 0.71212$ (all five figures), $e_1 = 0.882$, $c_1 = 0.934$ (fig. *a*), $e_1 = 0.88$, $c_1 = 0.935$ (fig. *b*), $e_1 = 0.878$, $c_1 = 0.936$ (fig. *c*). $e_1 = 0.87$, $c_1 = 0.94$ (fig. *d*), $e_1 = 0.868$, $c_1 = 0.941$ (fig. *e*).

make a transition from a system in which P^e is the smallest equilibrium cuckoo density to one in which P^c is the smallest, passing through $P^e = P^{ec} < P^c$, and $P^c = P^{ec} < P^e$, i.e., cases (6.10) and (6.11). At the first critical point we find a coexistence of three host types in stable equilibrium, namely all-acceptance, egg rejection and all-rejection. In fact, at this point there exists a one-parameter family of stable solutions. The realization of any steady state is dependent on initial conditions. A similar situation is found at the second critical point, illustrated in Figure 6.7*d*. Here we find a particular solution of the one-parameter family

of solutions in which all-accepters, chick-rejecters and all-rejecters are in coexistence.

Note that in Figure 6.7*a* and 6.7*e* we can identify the stable coexistence of two *defensive* host types as described in the previous section. For large values of k the all-rejecters are the only hosts present, despite the fact that $P^e < P^{ec}$ (fig.6.7*a*), or $P^c < P^{ec}$ (fig.6.7*e*) . We stress that this is not predicted by the heuristic argument.

Case (6.12) is illustrated in Figure 6.8. Also here we make a transition for a system in which egg-rejecters are fittest to one in which chick-rejecters are fittest, passing through the desired critical point. At the critical point we see that we again find coexistence of three host types, in this case all-accepters, egg- and chick-rejecters. Also here we find a one-parameter family of steady states at this critical point in parameter space. In this particular simulation we also find a Naimark-Sacker region in which steady oscillatory behaviour occurs confined between the two curves enclosing the region.

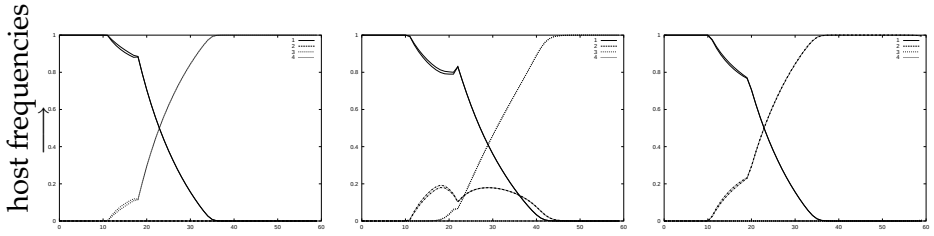


Figure 6.8: Here we have set P^{ec} at a large value to illustrate the effect of passing through a transition from egg rejection to chick rejection without encountering all-rejection. Now we find coexistence between all-accepters, egg- and chick-rejecters at the critical point as illustrated in Figure *b*. Moreover, we find Naimark-Sacker regions in all three pictures, in which quasi-periodic solutions are found. Parameters are as in Figure 6.2, but with the exception of $k_2 = 0.31$, $f = 0.71212$ (all three figures) and with $e_1 = 0.8754$, $c_1 = 0.9373$ (fig. a), $e_1 = 0.875$, $c_1 = 0.9375$ (fig. b), $e_1 = 0.8746$, $c_1 = 0.9377$ (fig. c). In this simulation we have not chosen k_2 conform its definition on page 174. However, the chosen value does comply with the assumption that $k_2 > e_2, c_2$, and as such can be seen as realistic.

We are left with the coincidence of all three cuckoo population equilibria, i.e., when $P^e = P^c = P^{ec}$. In this case, we find a two-parameter

family of solutions in which all four host types coexist in a stable steady state. The ratio in which the four host types coexist with each other depends on initial conditions. An example is given in Figure 6.9.

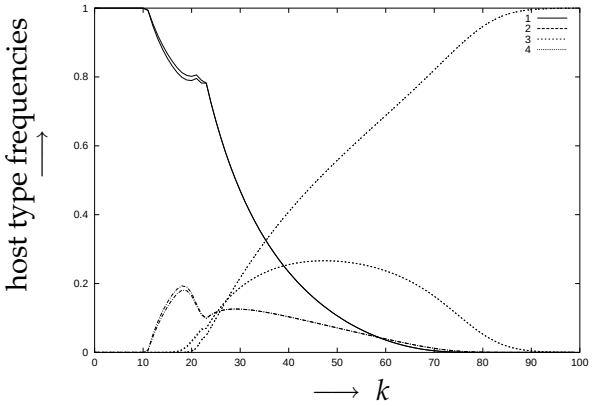


Figure 6.9: If all three equilibrium cuckoo densities are equal, the four host types coexist in intermediate frequencies. The actual ratios depend on initial conditions. Note the region with quasi-periodic solutions for k between 1000 and 2000 in the Naimark-Sacker-region. Curve (1) represents all-accepters, curve (2) egg-rejecters, curve (3) chick-rejecters and curve (4) all-rejecters. Parameters are as in Figure 6.2 except with $f = 0.71212$, $e_1 = 0.875$, $c_1 = 0.9375$, $k_1 = 0.83333$, $k_2 = 0.4$ (see Figure 6.8 for the choice of k_2).

6.3. Discussion and biological implications

With these results at our disposal, we will now give a reflection on some current views in ecology, and more specifically on the original question: why do we find so few cuckoo hosts that are able to defend themselves by recognizing an aberrant cuckoo chick in their nest?

Before giving a detailed account on the question of why chick rejection has not evolved among cuckoo hosts, we will first put the model in an ecological perspective.

Diploid model

We have also looked at a full diploid system (e.g., as in [133, 132]), and may conclude that although there are some quantitative differences, the overall effects of both models are qualitatively the same. The

host species	survival rate s_H
Reed Warbler (*Acrocephalus scirpaceus*)	$0.37 - 0.51$
Meadow Pipit (*Anthus pratensis*)	0.83
Wren (*Troglodytes troglodytes*)	$0.58 - 0.63$
Dunnock (*Prunella modularis*)	0.51
Redstart (*Phoenicurus phoenicurus*)	$0.51 - 0.71$

Table 1: Survival rates for various well-known host species of the cuckoo. From [30, 31].

main difference is to be seen in the invasion speeds. The process of recombination slows down the increase of the selectively advantageous hosts, and vice versa for selectively disadvantageous hosts.

Fitting the model to nature

As we have seen in the analysis, the model may display a sensitive dependence on parameters. For instance, we have found a small interval in the s_H-range, where sudden changes in the qualitative behaviour take place. We briefly compare this interval with experimental values. Table 1 shows the survival rates for a number of well-known hosts of the cuckoo.

Takasu *et al.* (1993) have used somewhat lower values ($s_H \sim 0.2, 0.5$). The s_H values in which the periodic solutions have been found are in the lower range of the experimental values. For instance, taking the parameter values of Figure 6.5 we find oscillatory behaviour roughly in the interval $s_H \in (0.365, 0.38)$ as shown in Figure 6.4. Hence we may have to be careful with the interpretation of population dynamics in species which exhibit survival rates at this lower end of the spectrum: they may display sustained defensive rejection behaviour for s_H values even below the critical asymptotic value.

Perfect defence is not sustainable

We have discussed the existence of the equilibria S^e, S^c and S^{ec} in which we find coexistence of two host types, namely the accepters and one defensive host type. In the Appendix we have given the full mathematical derivation of these solutions. At each of these three equilibria the values of the fitness functions of the respective host types are equal. This simplifies the procedure. To give an example, for S^e we find that egg-rejection hosts have the same fitness as accepter hosts at the given

cuckoo density. In mathematical terms, and recalling that

$$S^e = (P^e, H^e, h^a, h^e, 0, 0)$$

by equation (6.7), we have

$$f_a(P^e) = f_e(P^e).$$

From this expression we infer that this gives only relevant solutions provided $P^e > 0$, or

$$P^e = \frac{1}{a} \log \left(\frac{1 - e_1 + e_2}{e_2}\right) > 0.$$

Hence, if we were to exclude any discrimination costs by setting $p_e = 0$, and hence $e_1 = 1$, we would find $P^e = 0$. Similarly, if the chick-rejecters or all-rejecters were to defend themselves perfectly, giving cost coefficients of c_1 and $k_1 = 1$, we would obtain $P^c = 0$ and $P^{ec} = 0$ respectively. So for the cuckoo it is crucial that the hosts display imperfect defence behaviour. Also from a physiological point of view there may be consequences. Let us suppose that the defensive hosts have to incur some physiological cost associated with the ability to discriminate eggs or chicks (a cost we have so far neglected). Now, if we introduce a perfectly-defending host type in an otherwise accepting host population, the defending hosts will increase in number, and will drive the cuckoos to extinction. From that moment on, the two host types will be selectively neutral with respect to defence behaviour, but the accepting hosts will have a selective advantage in terms of the physiological costs they don't have to incur. Heuristically, we may thus conclude that in the long run we expect to find non-defending hosts again.

In short, the assumption that hosts birds defend themselves imperfectly by making type I errors (p_e, $p_c > 0$) is a necessary condition for any discrimination behaviour to be sustained within this model.

The non-prevalence of chick rejection

We now turn to more specific considerations and formulate more precisely why, according to our findings, chick rejection is rarely in nature.

Chick-rejection behaviour is incorporated in this model in two host types: the chick-rejecters and the all-rejecters. As we have seen, in most cases the heuristic argument given at the beginning of the Analysis section is applicable. Hence, to explain why we do not see chick-rejecters

rather than egg-rejecters we have to show that $P^e < P^c$ for realistic parameter values. The conditions for this are

$$\frac{1-e_1}{e_2} < \frac{1-c_1}{c_2}.$$

After some algebra this reduces to

$$\gamma < \gamma_{ec} = \frac{p_c}{1-q_c}\frac{1-q_e}{p_e}\frac{1-b_c}{1-b_e}.$$

Now $b_c \leq b_e$, so that $(1-b_c)/(1-b_e) \geq 1$, and γ, the average fraction of the clutch saved by parasitized chick-rejecters, satisfies $\gamma < 1$. The combination $p_e/(1-q_e)$ is a measure of how difficult it is to discriminate eggs. It increases with the probability of either type I or type II errors. Hence, unless it is more difficult by a sufficient margin to discriminate eggs than chicks, so that

$$\frac{p_e}{1-q_e}\frac{1-q_c}{p_c} \geq \frac{1-b_e}{1-b_c}\frac{1}{\gamma} > 1,$$

chick-rejection cannot compete with egg-rejection. If γ is small, as it might be unless chick-rejecters keep a close watch on their nest, then chick-discrimination must be much easier than egg-discrimination for chick-rejection to prevail. It has been suggested that to spot a cuckoo chick might be more difficult than to recognize a cuckoo egg (Davies and Brooke 1988). The chicks are born within a number of days and their appearance changes quickly due to rapid growth. Eggs may look more homogeneous. This suggests that $p_e/(1-q_e) \leq p_c/(1-q_c)$, and under these circumstances egg-rejecters always outcompete chick-rejecters.

All-rejecters are fitter than egg-rejecters if they are parasitized (since $k_2 > e_2$). To explain why we do not see all-rejecters rather than egg-rejecters we have to show that $P^e < P^{ec}$ for realistic parameter values. The conditions for this are

$$\frac{1-e_1}{e_2} < \frac{1-k_1}{k_2}.$$

Defining $\alpha = e_1 - k_1 > 0$, $\beta = k_2 - e_2 > 0$, this is equivalent to

$$\frac{1-e_1}{e_2} < \frac{\alpha}{\beta},$$

or

$$\gamma < \gamma_{ek} = \frac{1-b_c}{1-b_e}\frac{p_c}{1-q_c}\frac{1-q_e}{p_e}\frac{1-p_e}{q_e}\left\{1 + \frac{p_e(b_e - b_{ec})}{(1-p_e)(1-b_c)}\right\}.$$

The term in the braces is greater than 1. If egg-rejection errors are sufficiently small that $p_e + q_e < 1$, then $(1 - p_e)/q_e > 1$. Egg discrimination would be extremely poor if this inequality did not hold. Either p_e or q_e or both would be greater than $\frac{1}{2}$, meaning that the host would be more likely than not to make an error of type I, type II, or both. It follows that $\gamma_{ek} > \gamma_{ec}$, so that if the chick-rejecter inequality $\gamma < \gamma_{ec}$ holds, then

$$\gamma < \gamma_{ec} < \gamma_{ek},$$

and the all-rejecter inequality also holds. It is easier for chick-rejecters than all-rejecters to invade a steady state consisting of egg-rejecters and all-accepters. At first sight this is a surprising result, and it is interesting to see why it holds. We may trace it back to the fact that $\beta = k_2 - e_2$ is small unless errors are large. Even when all-rejecters have an advantage, which is when parasitism pressure is high, that advantage is small, only being brought into play when, through an error, the first line of defence has failed. Compounding the difficulty, the stable steady state, where the system ends up, is the one where parasitism pressure is lowest, and even this small advantage is likely to disappear.

This result may be applied much more generally. Similar arguments could be advanced for any situation where two consecutive lines of defence against parasitism or predation were possible, and we would predict that although in certain circumstances either one could prevail, the strategy of maintaining both defence systems would only be worthwhile if there was a high probability that the first one would fail.

Evolutionary lag or equilibrium?

There has been a long debate whether the observed lack of chick rejection (and also the variation in egg rejection in different host species) in avian brood parasitism is due to an evolutionary lag [38, 145, 87, 127] or whether the present situation is one of evolutionary equilibrium [34, 119]. We argue that the model captures both ideas, and by choosing discrimination costs appropriately, we will see that the model is in agreement with both of these views.

These hypotheses are both concerned with the explanation of one trait. If we want to view chick discrimination from the viewpoint of the lag hypothesis, we assume a lack of hereditary variation in the gene coding for this trait and sufficient benefit for the host to have the ability reject a cuckoo chick [117]. Hence we assume c_1 and c_2 to be large, so chick discrimination is quick (γ close to 1) and easy (p_c and q_c are close to zero). There are no restrictions on k_1, k_2, e_1 or e_2. The critical point

would be values of c_1 and c_2 such that

$$\frac{1-c_1}{c_2} = \frac{1-e_1}{e_2}. \tag{6.14}$$

The model predicts that, starting with any non-trivial number of chick-rejecters, if chick rejection is the attracting steady state, the system will converge to this steady state.

If we assume higher costs for chick-discrimination, so that it is either slow or difficult, hence assuming c_1 and c_2 to be low and thereby questioning the adaptiveness of this trait, we are in the domain of the equilibrium hypothesis. This view assumes moreover that there is sufficient hereditary variability in the host population, i.e., the number of chick-rejecters in the host population is non-trivial. Again there are no constraints on any of the other four discrimination costs. The critical point is again given by equation (6.14), and the model predicts that chick rejection should prevail if these discrimination costs are low enough.

The first supportive evidence for chick rejection

As we have discussed in the Introduction on page 170, there has been some debate whether or not hosts of parasitic birds need comparitive material to be able to discriminate their young from the parasite's. Two recent papers indicate that such simultaneous comparison is in fact not always necessary. In these cases the strategy used is different from the ones already suggested. Australian superb fairy-wrens desert all nests with a non-mimetic shining bronze-cuckoo chick present and 40% of nests with mimetic Horsfield's bronze-cuckoo chick present [81]. Fairy-wrens sometimes desert their own lone chicks but the desertion rate is significantly higher in the presence of the parasitic progeny. Thus, cuckoo chick desertion cannot be explained fully as a by-product of a life-history strategy to avoid inefficient parental investment in a single chick brood. Further experiments showed that the cue for recognition is not the appearance of the chick but the structure of begging calls. The host response was clearly not based on imprinting—females that accepted a parasitic chick did not abandon a lone host chick in later breeding attempts.

European reed warblers adopted similar responses to parasitic common cuckoo chicks—desertion after a very long nestling period [60]. Several lines of evidence indicated that warblers refused to feed cuckoo chicks that require a higher intensity of parental care than an

average host brood at fledging, i.e., feeding rates to the parasite were outside the normal range of parental care at an unparasitized nest.

How does the model fit the real world?

Almost all hosts of the common cuckoo show a relatively high rejection rate of parasitic eggs [36] which could—together with extremely low parasitism rates at the host species level ($< 5\%$ [33])—explain the absence of nestling discrimination in these hosts. Note that even without parasitic egg rejection by the host the effective parasitism rate at the chick stage is lower than at the egg stage because some parasitic eggs are ejected by another cuckoo (4%), and some are infertile, laid too late in the breeding cycle of the host, or do not hatch for some other reason (2%) ([105], Øien *et al.* unpublished data).

On the other hand, hosts of the brown-headed cowbird *Molothrus ater* are either close to 100% egg-rejecters (selection for nestling rejection is then nil) or they accept almost all parasite eggs. This is probably a result of a recent colonization by the parasite [119] so we cannot expect to observe nestling rejection here either. In all, both systems with evicting and non-evicting parasites are in line with the third prediction of the model that egg rejection is often the only defence strategy present within the host population.

Rejection of the eggs of parasites renders few opportunities that favour nestling recognition in comparison to egg recognition. Thus, the rare enemy hypothesis becomes the "rarer enemy hypothesis" which predicts that nestling discrimination (and consequently parasitic nestling mimicry) should predominantly evolve in hosts that display mild egg rejection, for instance by being tricked into accepting parasite eggs because of the nearly perfect match between parasitic and host eggs. Such a nearly perfect match could result from mimicry and phylogenetic [104] or physical constraints (a host cannot discriminate eggs in poor light conditions of dark domed nests especially if alien and host eggs are of similar size and shape [46]).

Several host-parasite systems that show 100% egg acceptance also discriminate against parasitic nestlings and parasite mimicry can also be found. Examples are estrildids parasitized by Vidua finches [104, 107], bay-winged cowbird *Molothrus badius* parasitized by screaming cowbird *M. rufoaxillaris* [46] and superb fairy-wrens parasitized by Horsfield's bronze-cuckoo [81]. Circumstantial evidence for joint occurrence of egg-acceptance and nestling rejection/mimicry comes from other two systems: house crows *Corvus splendens* parasitized by the

Asian koel *Eudynamys scolopacea* [39], and shining bronze-cuckoo parasitising the grey warbler *Gerygone igata* [57]. Both hosts accept all parasitic eggs. It has been argued that the chicks of Asian koel are probably mimetic [34, 33], and indirect evidence suggests also nestling mimicry in the shining bronze-cuckoo [57].

Hence, selection induced by parasitic nestlings is not relaxed by parasitic egg ejection in any of these systems. This is consistent with the rarer enemy hypothesis. It should be noted, that rufous-bellied thrushes *Turdus rufiventris* may also be able to discriminate against parasitic shiny cowbird *Molothrus bonariensis* chicks [85] but the accepter/rejecter status of this host is unfortunately unclear; this host is a very weak rejecter at best [84].

Some authors [111, 57] considered discrimination against foreign nestlings in a species that accept strange eggs to be puzzling. However, as egg rejection keeps the effective parasitism rate at the nestling stage at levels that might not allow positive selection of nestling discrimination, there are good reasons to expect exactly this pattern.

The model fits the two new papers reasonably well. The fairy-wrens adopt a chick rejection strategy that is both low cost and virtually error-free (fairy-wrens sometimes desert lone own chicks but one chick broods are extremely rare in this species). This fits the rarer enemy hypothesis as we have seen. Reed warblers have higher costs and probabilities of re-nesting are lower due to a shorter breeding period. Although there is substantial egg rejection ($\sim$40%) among the warblers [105], the relatively high parasitism rates in this particular population suggest it is not very effective. However, the warbler adopts a strategy of delayed chick rejection which could mean it is virtually certain that it has a cuckoo in the nest. Thus, it very rarely has to pay the costs of a recognition error - reed warblers were reported to never desert one-nestling broods [33, 59]. The advantage of error-free defence bears the cost of long care for the parasite before desertion. These factors could perhaps explain the lower frequency of chick desertion in warblers compared to fairy-wrens (16 vs 40-100%).

Theoretical implications and future directions

Following three decades in which brood parasitism research focused almost exclusively upon the egg stage the last five years have brought a set of exciting papers which finally pay attention to the nestling stage [60, 81, 46, 85, 107]. In future research we suggest that it would be best not to focus exclusively on parasite chick mimicry

because this may be distracting—chick discrimination by hosts may not necessarily lead to the evolution of nestling mimicry in parasites [60, 46]. A more fruitful approach would be to focus on (1) host responses (in species parasitized with near perfectly matching eggs) to cross-fostered non-parasitic nestlings of other species and (2) effects of brood reduction in both naive and experienced host individuals. The reason for the latter is that the costs of desertion may decrease with the age and experience of the host (rejection of parasitic nestlings by superb fairy-wrens is partly explained by brood reduction [81]).

This synthesis of current experimental and theoretical progress gives new insight into the interaction between defence strategies and the working of the rare enemy effect. It may also explain paradoxes in other parasitic systems. For example, when there is more than one level of parasitism, we can now understand why some species fail to recognize their own friends. Ant colonies *Myrmica schencki* are sometimes parasitized by butterfly larvae *Maculinea rebeli* [2]. In turn the butterfly larvae may be parasitized by *Ichneumon eumerus* wasps. A chemical cocktail that provokes in-fighting among the ant workers [134] allows the wasp to get access to and lay its egg in the caterpillar. Although the wasps help relieve the ants from their parasites (if the parasitized caterpillar consumes less of the ants' resources), the ants clearly do not recognize the wasps as their allies - or why would the wasp have to cause internecine warfare to get into their nest? We suggest that the beneficial wasp is simply too rare to be recognized as a friend. The wasp is rare because the butterfly is rare: the extreme specialization within this system drives a "rare friends effect".

In both cases, the actual rare enemy or rare friend is made rare by defences against other 'predators' in the system. Cuckoo chicks are rare because egg rejection has made them rare; the wasps are rare because the butterflies are. These two parasite-host systems elucidate subtle new mechanisms in which rarity structures the evolution of organisms, and indicate that the evolution of some adaptations could change the 'rules of the game' at other stages of coevolution, leading to maladaptiveness of some traits that would otherwise be adaptive for the bearers.

6.4. Appendix

We start the more formal mathematical investigations with an explicit formulation of the three equilibria S^e, S^c and S^{ec}. We will do these derivations for S^e and leave the rest of the details to the interested reader.

S^e is converged upon if the fitness of egg-rejecters and all-accepters is equal. By solving the fitness equations $f_a = f_e$ for P_t we find that at S^e

$$P_t = P^e \;:=\; \frac{1}{a}\log\Big(\frac{1-e_1+e_2}{e_2}\Big). \tag{6.15}$$

For easy reference in the discussion, the other equilibrium values for the cuckoo for S^c and S^{ec} are given here.

$$P_t = P^c \;:=\; \frac{1}{a}\log\Big(\frac{1-c_1+c_2}{c_2}\Big), \tag{6.16}$$

$$P_t = P^{ec} \;:=\; \frac{1}{a}\log\Big(\frac{1-k_1+k_2}{k_2}\Big). \tag{6.17}$$

If we introduce arbitrary (non-trivial) host type frequencies for all four host types, and the system converges to S^e, then we will see that the chick-rejecters and all-rejecters have vanished at steady state: they have been outcompeted by the fitter egg-rejecters. Hence, using eq. (6.15) and setting $h^c = h^{ec} = 0$, we can solve eq. (6.2) at steady state to find

$$H_t = H^e := k\big(s_H + \frac{fe_2}{1-e_1+e_2} - 1\big).$$

To express h^e at steady state we introduce

$$A(x) = \frac{(1-s_P)\log x}{aG(s_H + \frac{f}{x} - 1)(1-\frac{1}{x})}.$$

Now we find h^e and h^a at steady state by solving eq. (6.1):

$$\begin{aligned} h^a &= 1 - \frac{1}{1-q_e}\Big(1 - \frac{1}{k}A\Big(\frac{1-e_1+e_2}{e_2}\Big)\Big), \\ h^e &= \frac{1}{1-q_e}\Big(1 - \frac{1}{k}A\Big(\frac{1-e_1+e_2}{e_2}\Big)\Big). \end{aligned} \tag{6.18}$$

In conclusion, we find

$$S^e \;=\; (P^e, H^e, h^a, h^e, 0, 0).$$

The derivation of the other equilibria is done in a completely analygous fashion. For chick and all rejection the expressions are given upon substitution of c_i and k_i for e_i respectively, $i = 1, 2$, and similarly q_c and q_{ec} for q_e, in all equilibrium expressions for egg rejection. The formulae for the three other steady states—egg and all rejection, chick and all rejection and egg and chick rejection—are slightly more complicated but found in precisely the same way.

We make the following remarks, which apply for all three S^e, S^c and S^{ec} but to avoid iteration are only stated for S^e. The steady state S^e is illustrated in Figure 6.3 in terms of h^e with respect to our chosen bifurcation parameter k. It intersects $h^e = 0$ and $h^e = 1$ in two points. In the interval between these points we hence have coexistence of egg-rejecters with all-accepters. We extend S^e by including the steady states $h^e = 0$ and $h^e = 1$, and denote this by $\bar{S}^e$.

A necessary condition for S^e to be attained is $h^e > 0$. Using this condition, we can find a minimal value for k for this condition to be satisfied, denoted by k_0^e, such that $h^e > 0$:

$$k_0^e = A\Big(\frac{1 - e_1 + e_2}{e_2}\Big).$$

Recall that k^P is the critical value for the existence of the cuckoo population, given by $k^P = (1 - s_P)/aG(s_H + f - 1)$. Note that $k^P < k_0^e$. For k in the interval (k^P, k_0^e), the system converges to S^1.
Another necessary condition for S^e to be attained is that $h^e < 1$. Following the same procedure, we find a maximal value for k, denoted by k_1^e,

$$k_1^e = \frac{1}{q_e} A\Big(\frac{1 - e_1 + e_2}{e_2}\Big).$$

If we solve $h^e = 0$ for s_H we find an asymptotic value for which S^e can exist. Denoting it by s_H^e, it is given by

$$s_H^e = 1 - \frac{f e_2}{1 - e_1 + e_2}.$$

An illustration of this asymptote is given in Figure 6.4. Under the assumption that $e_1, c_1 \geq k_1$ and $k_2 \geq e_2, c_2$ one can show that $s_H^{ec} \geq s_H^e, s_H^c$, and $k_0^{ec} \geq k_0^e, k_0^c$.

Dynamical behaviour of the three equilibria

With these explicit equilibrium solutions we investigate which of these equilibria is attained by the system. Suppose that for instance $P^e < P^c < P^{ec}$, that k is sufficiently large, and that all host type frequencies are non-trivial. Then if we start at some $P_0 > P^{ec}$, we know by construction of the fitness functions, that the all-rejection hosts have highest fitness (see Figure 6.1). Hence, they will cause the cuckoo population to decrease. By the decreasing nature of f_e, there comes a point where egg rejection becomes fitter than all-rejection. This is reflected by the fact that we have assumed that $P^e < P^{ec}$. So the cuckoo population will decrease further and will eventually spiral around P^e: for $P_t < P^e$

the all-accepters are fittest and their increase allows P_t to increase; the opposite effect is seen when $P_t > P^e$. If we assume that these oscillations converge to an equilibrium, we only have one option: the system converges to S^e. This argument indicates that the relative position of the equilibrium cuckoo populations may determine to which solution the system converges.

As we have seen in the previous paragraph, we can determine to which equilibrium solution the system converges by looking at the relative positions of P^e, P^c and P^{ec}. It is immediate that $P^e < P^c$ if $\frac{1-e_1+e_2}{e_2} < \frac{1-c_1+c_2}{c_2}$, and analogous identities for the other options.

Bibliography

[1] P.K. Agarwal, H. Edelsbrunner, and Y. Wang. Computing the writhing number of a polygonal knot. In *13th ACM-SIAM Symp. Discrete Algorithms*, 2002.

[2] T. Akino, J.J. Knapp, J.A. Thomas, and G.W. Elmes. Chemical mimicry and host specificity in the butterfly *Maculinea rebeli*, a social parasite of *Myrmica* ant colonies. *Proc. Roy. Soc. London B*, 266:1419–1426, 1999.

[3] J. Aldinger, I. Klapper, and M. Tabor. Formulae for the calculation and estimation of writhe. *J. Knot Theory and its Ramifications*, 4(3):343–372, 1995.

[4] J.C. Alexander and S.S. Antman. The ambiguous twist of Love. *Quarterly of Applied Math.*, 40:83–92, 1982.

[5] R.M. Anderson and R.M. May. Population biology of infectious diseases. Part I. *Nature*, 280:361–367, 1979.

[6] J.B. André, J.B. Ferdy, and B. Godelle. Within-host parasite dynamics, emerging trade-off, and evolution of virulence with immune system. *Evolution*, 57(7):1489–1497, 2003.

[7] S.S. Antman. *Nonlinear problems in elasticity*. Springer-Verlag, 1995.

[8] D.K. Arrowsmith and C.H. Place. *Introduction to Dynamical Systems*. Cambridge University Press, 1990.

[9] J. Bemelmans and M. Chipot. On a variational problem for an elastic membrane supporting a heavy ball. *Calculus of Variations and PDE's*, 3:447–473, 1995.

[10] C. Benham. Geometry and mechanics of DNA superhelicity. *Biopolymers*, 22:2477–2495, 1983.

[11] J.G. Blom and M.A. Peletier. A continuum model of lipid bilayers. Technical Report MAS-R0229, CWI, 2002.

[12] C. Bouchiat and M. Mézard. Elasticity model of a supercoiled DNA molecule. *Phys. Rev. Lett.*, 80:1556, 1998.

[13] J. Bourgain, H. Brezis, and P. Mironescu. Another look at Sobolev spaces. In J.L. Menardi, E. Rofman, and A. Sulem, editors, *Optimal Control and Partial Differential Equations*, a volume in honor of A. Bensoussans's 60th birthday, pages 439–455. IOS Press, 2001.

[14] H. Brezis. *Analyse fonctionelle*. Masson, Paris, 1983.

[15] H. Brezis and D. Kinderlehrer. The smoothness of solutions to nonlinear variational inequalities. *Indiana Univ. Math. J.*, 23(9):831–844, 1974.

[16] L.C. Brooker, M.G. Brooker, and A.M.H. Brooker. An alternative population/genetic model for the evolution of egg mimesis, and egg crypsis in cuckoos. *J. Theor. Biol.*, 146:123–143, 1990.

[17] A. Burchard. Cases of equality in the Riesz rearrangement inequality. *Ann. of Math. (2)*, 143(3):499–527, 1996.
[18] L.A. Caffarelli. Free boundary problem in higher dimensions. *Acta Math.*, 139:155–184, 1977.
[19] L.A. Caffarelli. Compactness methods in free boundary problems. *Comm. Part. Diff. Eq.*, 5(4):427–448, 1980.
[20] L.A. Caffarelli. A remark on the Hausdorff measure of a free boundary, and the convergence of coincidence sets. *Boll. Un. Mat. Ital. A*, 18(5):109–113, 1981.
[21] L.A. Caffarelli. The obstacle problem revisited. *J. Fourier Anal. Appl.*, 4:383–402, 1998.
[22] G. Călugăreanu. Sur les classes d'isotopie des nœuds tridimensionnels et leurs invariants. *Czechoslovak Math. J.*, 11:588–625, 1961.
[23] J. Cantarella. On comparing the writhe of a smooth curve to the writhe of an inscribed polygon. To appear in SIAM Journal of Numerical Analysis. (arXiv: math.DG/0202236), 2002.
[24] J. Cantarella, J.H.G. Fu, R. Kusner, J.M. Sullivan, and N.C. Wrinkle. Criticality for the Gehring link problem. arXiv: math.DG/0402212, 2004.
[25] J. Cantarella, R.B. Kusner, and J.M. Sullivan. On the minimum rope length of knots and links. *Inventiones Mathematicae*, 150:257–286, 2002.
[26] A.R. Champneys, G.W. Hunt, and J.M.T. Thompson. *Localization and Solitary Waves in Solid Mechanics*. Advanced Series in Nonlinear Dynamics, Vol. 12. World Scientific, Singapore, 1999.
[27] B.D. Coleman and D. Swigon. Theory of supercoiled elastic rings with self-contact and its application to DNA plasmids. *J. Elasticity*, 60:173–221, 2000.
[28] B.D. Coleman, D. Swigon, and I. Tobias. Elastic stability of DNA configurations II. Supercoiled plasmids with self-contact. *Physical Review E*, 61:759–770, 2000.
[29] S.C. Cotter, L.E.B. Kruuk, and K. Wilson. Costs of resistance: genetic correlations and potential trade-offs in an insect immune system. *J. Evol. Biol.*, 17(2):421–429, 2004.
[30] S. Cramp. *Handbook of the birds of Europe, the Middle East and North Africa*, volume V. Oxford University Press, New York, 1988.
[31] S. Cramp and D.J. Brooks. *Handbook of the birds of Europe, the Middle East and North Africa*, volume VI. Oxford University Press, New York, 1992.
[32] B. Dacorogna. *Direct methods in the calculus of variations*. Springer-Verlag, 1989.
[33] N.B. Davies. *Cuckoos, cowbirds and other cheats*. T. & A.D. Poyser, London, 2000.
[34] N.B. Davies and M. de L. Brooke. Cuckoos versus reed warblers: adaptations and counteradaptations. *Anim. Beh.*, 36:262–284, 1988.
[35] N.B. Davies and M. de L. Brooke. An experimental study of co-evolution between the cuckoo, *Cuculus canorus*, and its hosts. II. host egg markings, chick discrimination and genera, discussion. *J. Anim. Ecology*, 58:225–236, 1989.
[36] N.B. Davies and M. de L. Brooke. An experimental study of co-evolution between the cuckoo, *Cuculus canorus*, and its hosts. I. host egg discrimination. *J. Anim. Ecology*, 58:207–224, 1989a.
[37] N.B. Davies, R.M. Kilner, and D.G. Noble. Nestling cuckoos *Cuculus canorus* exploit hosts with begging calls that mimic a brood. *Proc. Roy. Soc. London B*, 265:673–678, 1998.
[38] R. Dawkins and J.R. Krebs. Arms races between and within species. *Proc. Roy. Soc. London B*, 205:489–511, 1979.

[39] D. Dewar. An inquiry into the parasitic habits of the indian koel. *J. Bombay Nat. Hist. Soc.*, 17:765–782, 1907.

[40] E. Doedel, A. Champneys, T. Fairgrieve, Y. Kuznetsov, B. Sandstede, and X. Wang. Auto97: Continuation and bifurcation software for ordinary differential equations; available by ftp from ftp.cs.concordia.ca in directory pub/doedel/auto.

[41] S. Eis, E.H. Garman, and L.F. Ebel. Relation between cone production and diameter increment of douglas fir (*Pseudotsuga menziesii* (Mirb). Franco), grand fir (*Abies grandis* Dougl.) and western white pine (*Pinus monticola* Dougl.). *Canadian Journal of Botany*, 43:1553–1559, 1965.

[42] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions*. CRC Press, Boca Raton, 1991.

[43] B. Fain, J. Rudnik, and S. Östlund. Conformations of linear DNA. *Phys. Rev. E*, 55(6):7364–7368, 1997.

[44] R.P. Feynman and S. Weinberg. *Elementary Particles and the Laws of Physics: The 1986 Dirac Memorial Lectures*. Cambridge University Press, 1987.

[45] J.G.E.M. Fraaije, B.A.C. van Vlimmeren, N.M. Maurits, M. Postma, O.A. Evers, C. Hoffmann, P. Altevogt, and G. Goldbeck-Wood. The dynamic mean-field density functional method and its application to the mesoscopic dynamics of quenched block copolymer melts. *J. Chem. Phys.*, 106(10):4260–4269, 1997.

[46] R.M. Fraga. Interactions of the parasitic screaming and shiny cowbirds (*Molothrus rufoaxillaris* and *M. bonariensis*) with a shared host, the bay-winged cowbird (*M. badius*). In S.I. Rothstein and S.K. Robinson, editors, *Parasitic birds and their hosts*. Oxford University Press, 1998.

[47] W.B. Fraser and D.M. Stump. The equilibrium of the convergence point in two-strand yarn plying. *Int. J. Solids Structures*, 35(3–4):285–298, 1998.

[48] J. Frehse. On the regularity of the solution of a second order variational inequality. *Boll. Un. Mat. Ital. B*, 6(4):312–315, 1972.

[49] A. Friedman. *Variational inequalities and free-boundary problems*. John Wiley & Sons, New York, 1982.

[50] F.B. Fuller. The writhing number of a space curve. *Proc. Nat. Acad. Sciences USA*, 68(4):815–819, 1971.

[51] F.B. Fuller. Decomposition of the linking number of a closed ribbon: a problem from molecular biology. *Proc. Nat. Acad. Sciences USA*, 75:3557–3561, 1978.

[52] D.J. Futuyma. *Evolutionary Biology*. Sinauer Associates, Sunderland, Massachusetts, 1979.

[53] Zs. Gáspár and R. Németh. Shape finding of an extremely twisted ring. to appear in J. of Comp. and Appl. Mech., 2004.

[54] W.M. Gelbart, A. Ben-Shaul, and D. Roux, editors. *Micelles, Membranes, Microemulsions and Monolayers*. Springer-Verlag, 1995.

[55] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*. Springer-Verlag, 1977.

[56] L.E. Gilbert. Food web organization and the conservation of neotropical diversity. In M.E. Soulé and B.A. Wilcox, editors, *Conservation Biology*, pages 11–33. Sinauer Associates, Sunderland Massachusetts, 1980.

[57] B.J. Gill. Behavior and ecology of the shining cuckoo, *Chrysococcyx lucidus*. In S.I. Rothstein and S.K. Robinson, editors, *Parasitic birds and their hosts*, pages 143–151. Oxford University Press, 1998.

[58] O. Gonzalez, J.H. Maddocks, F. Schuricht, and H. von der Mosel. Global curvature and self-contact of nonlinearly elastic curves and rods. *Calculus Variations*, 14:29–68, 2002.

[59] T. Grim and M. Honza. Does supernormal stimulus influence parental behaviour of the cuckoo's host? *Behav. Ecol. Sociobiol.*, 49:322–329, 2001.

[60] T. Grim, O. Kleven, and O. Mikulica. Nestling discrimination without recognition: a possible defence mechanism for hosts towards cuckoo parasitism? *Proc. Roy Soc. London B (Suppl.)*, 270:S73–S75, 2003.

[61] T. Grim, R. Planqué, N.F. Britton, and N.R. Franks. May defence strategy antagonisms and rare enemy effects explain the rarity of parasitic nestling discrimination? submitted to J. Avian Biology, 2004.

[62] J.H. Hannay. Cyclic rotations, contractibility and Gauss-Bonnet. *J. Phys. A*, 31:L321–L324, 1998.

[63] M.E. Hauber and P.W. Sherman. Self-referent phenotype matching: theoretical considerations and empirical evidence. *Trends Neuroscience*, 24:609–616, 2001.

[64] J.W.S. Hearle and A.E. Yegin. The snarling of highly twisted monofilaments part II: cylindrical snarling. *Journal of the Textile Industry*, 63(9):490–501, 1972.

[65] G.H.M. van der Heijden. The static deformation of a twisted elastic rod constrained to lie on a cylinder. *Proc. Roy. Soc. London A*, 457:695–715, 2001.

[66] G.H.M. van der Heijden, S. Neukirch, V.G.A. Goss, and J.M.T. Thompson. Instability and self-contact phenomena in the writhing of clamped rods. *Int. J. Mechanical Sciences*, 45:161–196, 2003.

[67] G.H.M. van der Heijden, M.A. Peletier, and R. Planqué. A consistent treatment of link and writhe for open rods. submitted to Arch. Rat. Mech. Anal. arXiv: math-ph/0310057, 2003.

[68] G.H.M. van der Heijden, M.A. Peletier, and R. Planqué. Self-contact for rods on cylinders. submitted to Arch. Rat. Mech. Anal. arXiv: math-ph/0411007, 2004.

[69] G.H.M. van der Heijden and J.M.T. Thompson. Lock-on to tape-like behaviour in the torsional buckling of anisotropic rods. *Physica D*, 112:201–224, 1998.

[70] G.H.M. van der Heijden and J.M.T. Thompson. Helical and localised buckling in twisted rods: a unified analysis of the symmetric case. *Nonlinear Dynamics*, 21:71–99, 2000.

[71] P. Holmes, G. Domokos, J. Schmitt, and I. Szeberényi. Constrained Euler buckling: an interplay of computation and analysis. *Comput. Methods Appl. Mech. Engrg.*, 170:175–207, 1999.

[72] J.N. Israelachvili and G.E. Adams. Direct measurements of long-range forces between two mica surfaces in 10-4 to 1 M KNO3 solutions. *Nature*, 262:774–776, 1976.

[73] T.S. Johnson and M. Zuk. Parasites and tradeoffs in the immune response of female red jungle fowl. *Oikos*, 86(3):487–492, 1999.

[74] L. Kaufmann. *Knots and physics*. World Scientific, 1991.

[75] B. Kawohl. When are solutions to nonlinear elliptic boundary value problems convex? *Comm. Part. Diff. Eq.*, 10:1213–1225, 1985.

[76] C. Kelly. A model to explore the rate of spread of mimicry and rejection in hypothetical populations of cuckoos and their hosts. *J. Theor. Biol.*, 125:283–299, 1987.

[77] D. Kinderlehrer and L. Nirenberg. Regularity in free boundary problems. *Ann. Scuola Norm. Sup. Pisa*, 4:373–391, 1977.

[78] K. Klenin and J. Langowski. Computation of writhe in modeling of supercoiled DNA. *Biopolymers*, 54:307–317, 2000.
[79] D. Lack. *Ecological adaptations for breeding in birds*. Muethen, 1968.
[80] D.C. Lahti and A.R. Lahti. How precise is egg discrimination in weaverbirds? *Anim. Behav.*, 63:1135–1142, 2002.
[81] N.E. Langmore, S. Hunt, and R.M. Kilner. Escalation of a coevolutionary arms race through host rejection of brood parasitic young. *Nature*, 422:157–160, March 2003.
[82] M.J. Lawes and T.R. Marthews. When will rejection of parasite nestlings by hosts of nonevicting avian brood parasites be favored? A misimprinting-equilibrium model. *Behav. Ecol.*, 14(6):757–770, 2003.
[83] P. Lévy. *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, 1954.
[84] G. Lichtenstein. Parasitism by shiny cowbirds of rufous-bellied thrushes. *Condor*, 100:680–687, 1998.
[85] G. Lichtenstein. Low success of shiny cowbird chicks parasitizing rufous-bellied thrushes: chick-chick competition or parental discrimination? *Anim. Behav.*, 61:401–413, 2001.
[86] P.-L. Lions. The concentration-compactness principle in the Calculus of Variations. The locally compact case. I. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 1(2):109–145, 1984.
[87] A. Lotem. Learning to recognize nestlings is maladaptive for cuckoo *Cuculus canorus* hosts. *Nature*, 362:743–745, 1993.
[88] A. Lotem, H. Nakamura, and A. Zahavi. Rejection of cuckoo eggs in relation to host age—a possible evolutionary equilbrium. *Behav. Ecol.*, 3:128–132, 1992.
[89] A. Lotem, H. Nakamura, and A. Zahavi. Constraints on egg discrimination and cuckoo-host co-evolution. *Anim. Behav.*, 49:1185–1209, 1995.
[90] A.E.H. Love. *A treatise on the mathematical theory of elasticity*. Dover Publications, 4th edition, 1944.
[91] D.G. Luenberger. *Optimization by vector space methods*. John WIley & Sons, New York, 1969.
[92] J.H. Maddocks. Stability and folds. *Arch. Rat. Mech. Anal.*, 99:301–328, 1987.
[93] A.C. Maggs. Writhing geometry at finite temperature: random walks and geometric phases for stiff polymers. *J. Chem. Phys.*, 114:5888–5896, 2001.
[94] K. Marchetti. Costs to defence and the persistence of parasitic cuckoos. *Proc. Roy. Soc. London B*, 248:41–45, 1992.
[95] J.F. Marko. Supercoiled and braided DNA under tension. *Phys. Rev. E*, 55(2):1758–1772, 1997.
[96] R.M. May and R.M. Anderson. Population biology of infectious diseases. Part II. *Nature*, 280:455–461, 1979.
[97] R.M. May and S.K. Robinson. The population dynamics of avian brood parasitism. *Am. Nat.*, 126:475–494, 1985.
[98] A. Moksnes, E. Røskaft, and A.T. Braa. Rejection behavior by common cuckoo hosts towards artificial brood parasite eggs. *Auk*, 108:348–354, 1991.
[99] J.D. Moroz and P. Nelson. Entropic elasticity of twist-storing polymers. *Macromolecules*, 31:6333–6347, 1998.
[100] H. Nakamura. Brood parasitism by the cuckoo *Cuculus canorus* in japan and the start of new parasitism on the azure-winged magpie *Cyanopica cyana*. *Japanese J. of Orn.*, 39:1–18, 1990.

[101] S. Neukirch and G.H.M. van der Heijden. Geometry and mechanics of uniform n-plies: from engineering ropes to biological filaments. *J. Elasticity*, 69:41–72, 2002.
[102] S. Neukirch and G.H.M. van der Heijden. Geometry and mechanics of uniform n-plies: from engineering ropes to biological filaments. *Journal of Elasticity*, 69:41–72, 2002.
[103] A.J. Nicholson and V.A. Bailey. The balance of animal populations, I. *Proc. Zool. Soc. London*, 1:551–598, 1935.
[104] J. Nicolai. Der brutparasitismus der viduinae als ethologisches problem. *Z. Tierpsychol*, 21:129–204, 1964.
[105] I.J. Øien, A. Moksnes, E. Røskaft, and M. Honza. Costs of cuckoo cuculus canorus parasitism to reed warblers acrocephalus scirpaceus. *J. Avian Biol.*, 29:209–215, 1998.
[106] E. Orlandini, M.C. Test, S.G. Whittington, D.W. Sumners, and E.J. Janse van Rensburg. The writhe of a self-avoiding walk. *J. Phys. A: Math. Gen.*, 27:L333–L338., 1994.
[107] R.B. Payne, J.L. Woods, and L.L. Payne. Parental care in estrildid finches: experimental tests of a model of vidua brood parasitism. *Anim. Behav.*, 62:473–483, 2001.
[108] M.A. Peletier. Sequential buckling: A variational analysis. *SIAM J. Math. Anal.*, 32:1142–1168, 2001.
[109] R. Planqué, N.F. Britton, N.R. Franks, and M.A. Peletier. The adaptiveness of defence strategies against cuckoo parasitism. *Bull. Math. Biol.*, 64(6):1045–1068, 2002.
[110] M.H. Protter and H.F. Weinberger. *Maximum principles in differential equations*. Prentice-Hall, 1967.
[111] T. Redondo. Exploitation of host mechanisms for parental care by avian brood parasites. *Etología*, 3:235–297, 1993.
[112] F. Riesz. Sur une inégalité intégrale. *J. London Math. Soc.*, 5:162–168, 1930.
[113] S. Rohwer, C.D. Spaw, and E. Røskaft. Costs to northern orioles of puncture-ejecting parasitic cowbird eggs from their nests. *Auk*, 106:734–738, 1989.
[114] D. Rolfsen. *Knots and Links*. Publish or Perish, Berkeley, California, 1976.
[115] E. Røskaft, G.H. Orians, and L.D. Beletsky. Why do red-winged blackbirds accept eggs of brown-headed cowbirds? *Evol. Ecol.*, 4:35–42, 1990.
[116] V. Rossetto and A.C. Maggs. Writhing geometry of open DNA. *J. Chem. Phys.*, 118:9864–9874, 2003.
[117] I. Rothstein. Evolutionary rates and host defenses against avian brood parasitism. *Am. Nat.*, 109:161–176, 1975.
[118] S.I. Rothstein. Mechanisms of avian egg recognition: possible learned and innate factors. *Auk*, 91:796–807, 1974.
[119] S.I. Rothstein. A model system for coevolution: avian brood parasitism. *Ann. Rev. Ecol. Syst.*, 21:481–508, 1990.
[120] S.I. Rothstein. Relic behaviours, coevolution and the retention versus loss of host defences after episodes of avian brood parasitism. *Anim. Behav.*, 61:95–107, 2001.
[121] S.I. Rothstein and S.K. Robinson. Major unresolved questions in the study of avian brood parasitism. In S.I. Rothstein and S.K. Robinson, editors, *Brood-parasites and their hosts*, pages 419–425. Oxford University Press, 1998.
[122] T. Schlick. Modeling superhelical DNA: Recent analytical and dynamic approaches. *Curr. Opinion Struct. Biol.*, 5:245–262, 1995.

[123] F. Schuricht and H. von der Mosel. Characterization of ideal knots. *Calculus Variations*, 19:281–315, 2004.

[124] F. Schuricht and H. von der Mosel. Euler-Lagrange equations for nonlinearly elastic rods with self-contact. *Arch. Rat. Mech. Anal.*, 168:35–82, 2003.

[125] T. Slagsvold, B.T. Hansen, L.E. Johannessen, and J.T. Lifjeld. Mate choice and imprinting in birds studied by cross-fostering in the wild. *Proc. Roy. Soc. London B*, 269:1449–1455, 2002.

[126] S.B. Smith, L. Finzi, and C. Bustamente. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science*, 258:1122–1126, 1992.

[127] M. Soler, J.J. Soler, J.G. Martinez, and A.P. Møller. Chick recognition and acceptance: a weakness in magpies exploited by the parasitic great spotted cuckoo. *Behav. Ecol. Sociobiol.*, 37:243–248, 1995.

[128] E.L. Starostin. Comment of 'Cyclic rotations, contractibility and Gauss-Bonnet'. *J. Physics A: Math. Gen.*, 35:6183–6190, 2002.

[129] E.L. Starostin. On the writhe of non-closed curves. arXiv.org:physics/0212095, 2002.

[130] T.R. Strick, J.F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. The elasticity of a single supercoiled DNA molecule. *Science*, 271:1835–1837, 1996.

[131] D.M. Stump and G.H.M. van der Heijden. Birdcaging and the collapse of rods and cables in fixed-grip compression. *Int. J. Solids and Structures*, 38:4265–4278, 2001.

[132] F. Takasu. Why do all host species not show defense against avian brood parasitism: evolutionary lag or equilibrium? *Am. Nat.*, 151:193–205, 1998.

[133] F. Takasu, K. Kawasaki, H. Nakamura, J.E. Cohen, and N. Shigesada. Modeling the population dynamics of a cuckoo-host association and the evolution of host defences. *Am. Nat.*, 142:819–839, 1993.

[134] J.A. Thomas, J.J. Knapp, T. Akino, S. Gerty, S. Wakamura, D.J. Simcox, J.C. Wardlaw, and G.W. Elmes. Parasitoid secretions provoke ant warfare. *Nature*, 417:505–506, May 2002.

[135] J.M.T. Thompson, G.H.M. van der Heijden, and S. Neukirch. Supercoiling of DNA plasmids: mechanics of the generalized ply. *Proc. R. Soc. London, Series A*, 458:950–985, 2002.

[136] I. Tobias, D. Swigon, and B.D. Coleman. Elastic stability of DNA configurations I. general theory. *Physical Review E*, 61:747–758, 2000.

[137] D.L. Venable. Size-number tradeoffs and the variation of seed size with plant resource status. *Am. Nat.*, 140:287–304, 1992.

[138] J.K. Victoria. Clutch characteristics and egg discriminative ability of the african village weaver (ploceus cucullatus). *Ibis*, 114:367–376, 1972.

[139] E.A. Vogler. Structure and reactivity of water at biomaterial surfaces. *Advances in Colloid and Interface Science*, 74:69–117, 1998.

[140] A.V. Vologodskii and J.F. Marko. Extension of torsionally stressed DNA by external force. *Biophys. J.*, 73:123–132, 1997.

[141] G.S. Weiss. A homogeneity improvement approach to the obstacle problem. *Invent. Math.*, 138:23–50, 1999.

[142] J.H. White. Self-linking and the Gauss integral in higher dimensions. *Amer. J. Math.*, 91:693–728, 1969.

[143] R. Winfree. Cuckoos, cowbirds and the persistence of brood parasitism. *Trends Ecol. Evol.*, 14:338–343, 1999.
[144] A.C. Zaanen. *Continuity, Integration and Fourier Theory*. Springer-Verlag, Berlin Heidelberg, 1989.
[145] A. Zahavi. Parasitism and nest predation. *Am. Nat.*, 113:157–159, 1979.

Samenvatting

In dit proefschrift worden verschillende problemen onderzocht uit de toegepaste wiskunde. In elk van deze staat de rol van restricties centraal. In hoofdstukken 2 tot en met 5 worden deze beperkingen vooraf opgelegd; in het laatste hoofdstuk vinden we binnen het bestudeerde model achteraf een restrictie die een nieuw inzicht geeft in de behandelde problematiek.

Hoofstuk 1 geeft een inleiding in de verscheidene problemen die behandeld worden in dit proefschrift, en geeft kort weer welke resultaten worden besproken in de erop volgende hoofdstukken.

Hoofdstuk 2 heeft als onderwerp hoe een dun elastisch snoer kronkelt rond een cylinder. Tijdens het kronkelen kan het snoer met zichzelf in aanraking komen. Dit compliceert de wiskundige analyse, wat de inspiratie vormt van dit stukje onderzoek. We leiden een energieminimalizatie-probleem af met een niet-lokale restrictie, en geven een volledige karakterisatie van de oplossingen van dit systeem. Het voornaamste resultaat is een bewijs dat de verzameling punten waarin het snoer zichzelf aanraakt een gesloten interval vormt.

Hoofstuk 3 handelt over de extensie van drie bekende concepten uit de studie van gesloten snoeren (waarbij de einden aan elkaar vast zitten), Link, Twist en Writhe, naar de klasse van snoeren die voor onze toepassingen interessant zijn: open snoeren. De voornaamste uitdaging is te bestuderen voor welke deelverzameling open snoeren zulke nieuwe concepten gedefinieerd kunnen worden, zonder dat hierbij ambiguïteit ontstaan, en zonder verlies van de oorspronkelijke eigenschappen.

Hoofdstukken 4 en 5 behandelen een eenvoudig model voor de beschrijving van de vorming van membraan-achtige structuren die onstaan door de interactie tussen lipide moleculen en het omringende water waarin deze zijn opgelost. Nadat het model is geïntroduceert bekijken we de eerst de regulariteit van oplossingen. Vervolgens bestuderen we, gewapend met deze informatie, voor welke parameterwaarden in het model oplossingen bestaan.

Hoofdstuk 6 vormt het sluitstuk van dit proefschrift. Hier bekijken we het parasitaire gedrag van de Koekoek op de waardvogels bij wie zij haar eieren legt. Met name de verschillende strategieën waarmee de gastheren zichzelf kunnen verdedigen staan centraal. We laten zien dat het bijna nooit adaptief is voor de gastheer om zich te verdedigen met meer dan één strategie.

Dit proefschrift is gebaseerd op de volgende artikelen:

- G.H.M. van der Heijden, M.A. Peletier, and R. Planqué, *Self-contact for rods on cylinders*, submitted to Arch. Rat. Mech. Anal. arXiv: math-ph/0411007, 2004. (Hoofdstuk 2)
- G.H.M. van der Heijden, M.A. Peletier, and R. Planqué, *A consistent treatment of link and writhe for open rods*, submitted to Arch. Rat. Mech. Anal. arXiv: math-ph/0310057, 2003. (Hoofdstuk 3)
- M.A. Peletier and R. Planqué, *A simple continuum model for lipid membranes*, in preparation. (Hoofdstukken 4 and 5)
- R. Planqué, N.F. Britton, N.R. Franks, and M.A. Peletier, *The adaptiveness of defence strategies against cuckoo parasitism*, Bull. Math. Biol. **64** (2002), no. 6, 1045–1068. (Hoofdstuk 6)
- T. Grim, R. Planqué, N.F. Britton, and N.R. Franks, *May defence strategy antagonisms and rare enemy effects explain the rarity of parasitic nestling discrimination?*, submitted to J. Avian Biology, 2004. (Hoofdstuk 6)

Curriculum Vitae

De auteur van dit proefschrift werd geboren op 12 April 1977 te Leiderdorp. Hij behaalde het VWO-diploma aan het Stedelijk Gymnasium te Leiden in 1995. Hij zette zijn opleiding voort met de studie Wiskunde aan de Universiteit Leiden, waarbij hij aan het eind als Erasmus-student een half jaar in Bath, Groot-Brittannië verbleef. De studie werd in 2000 afgesloten.

Het promotieonderzoek verrichtte hij als onderzoeker in opleiding aan het Centrum voor Wiskunde en Informatica te Amsterdam, van 1 december 2000 tot 31 december 2004. Het vormde een onderdeel van het onderzoeksproject 'PDEs in the Life Sciences', en werd afgesloten met dit proefschrift. Het onderzoek stond onder leiding van prof.dr. M.A. Peletier.